FIGURE 1

Cloning Scheme for Generating pCMV-NS35
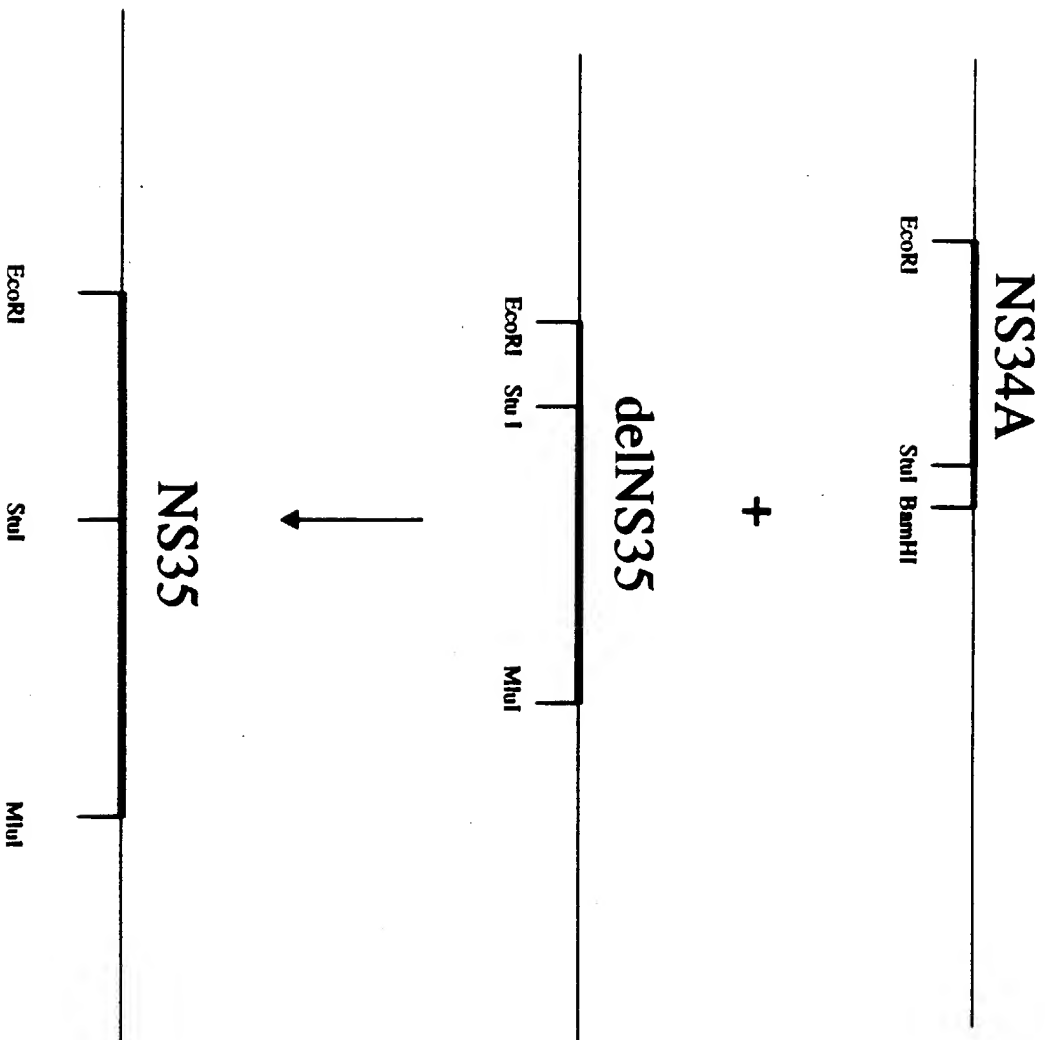
NS34A

EcoRI   StuI   BamHI

+

delNS35

EcoRI   StuI   MluI

NS35

EcoRI   StuI   MluI

FIGURE 2

Stu I (211)

EcoR I (1983)

pCMV-NS35
9621 bp

Stu I (2962)

Mlu I (7382)
BamH I (7373)

EcoR I (6135)

BamH I (5459)

BamH I (4285)
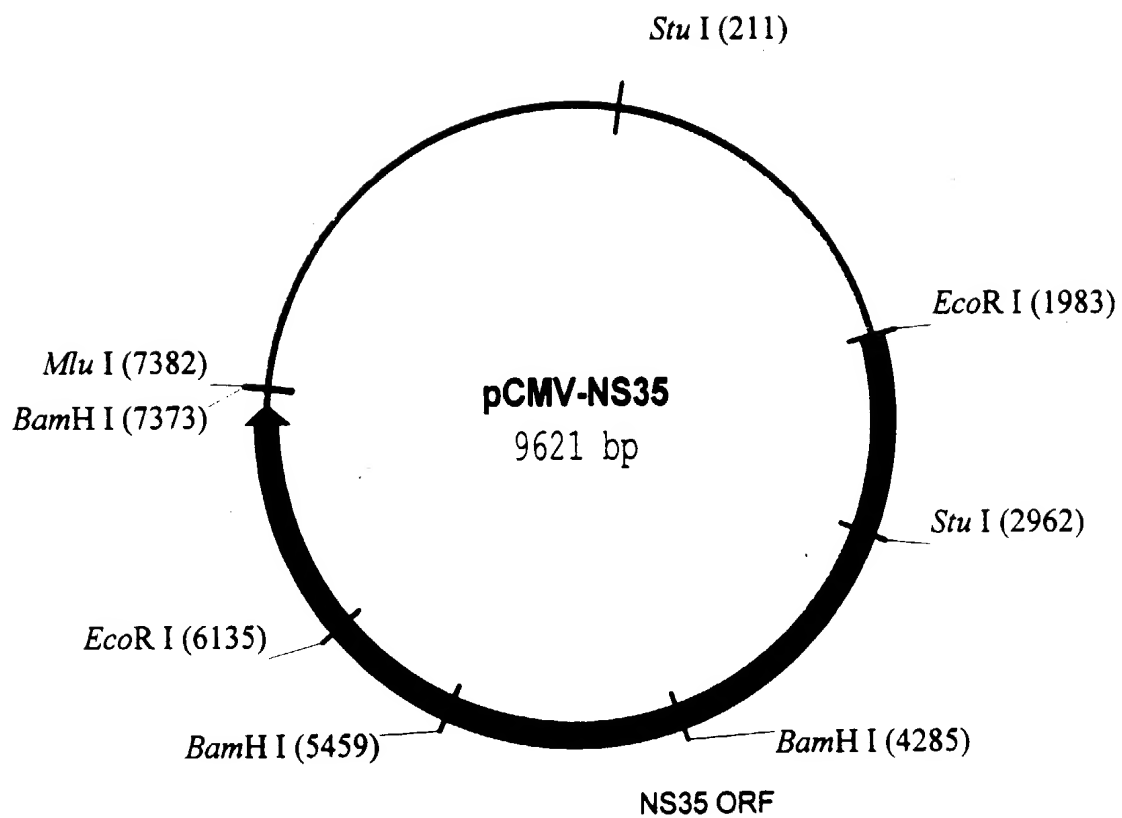
NS35 ORF

FIGURE 3 - Page 1

```
   1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT
      AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTCGCCTA

  81  GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA
      CGGCCCTCGT CTGTTCGGGC AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC GCCGTAGTCT

                                                          StuI
                                                        ~~~~~~~
 161  GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG
      CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC

 241  AATAGCTCAG AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGGCGGA
      TTATCGAGTC TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT ACCCCGCCTC TTACCCGCCT

 321  ACTGGGCGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
      TGACCCGCCC CTCCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

 401  CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT
      GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA

 481  AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCCATT
      TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA

 561  GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT
      CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

 641  AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG TCAATGACGG TAAATGGCCC
      TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC AGTTACTGCC ATTTACCGGG

 721  GCCTGGCATT ATGCCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
      CGGACCGTAA TACGGGTCAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

 801  CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA
      GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT

 881  TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC CCCGTTGACG
      AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

 961  CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCC CCTGGAGACG
      GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1041  CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG GGAACGGTGC ATTGGAACGC
      GGTAGGTGCG ACAAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCCGGC CCTTGCCACG TAACCTTGCG

1121  GGATTCCCCG TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
      CCTAAGGGGC ACGGTTCTCA CTGCATTCAT GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201  CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA
      GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

1281  TTGACCATTA TTGACCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACTAT
      AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

1361  CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT
      GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA
```

```
1441  TATTTACAAA TTCACATATA CAACAACGCC GTCCCCGTG CCCGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT
      ATAAATGTTT AAGTGTATAT GTTGTTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT ATCGCACCCT AGAGGCTGTA

1521  CTCGGGTACG TGTTCCGGAC ATGGGCTCTT CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
      GAGCCCATGC ACAAGGCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601  GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC
      CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTTACGG GTGGTGGTGG

1681  AGTGTGCCGC ACAAGGCCGT GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT
      TCACACGGCG TGTTCCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA CCTGCGTCTA

1761  GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT
      CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841  TGCGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT
      ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA
```

```
      +2                                                                        M   A   A
                                                                 EcoRI
                                                                 ~~~~~~
1921  GACAGACTAA CAGACTGTTC CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA GAATTCACC ATGGCTGCAT
      CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT CTTAAGTGG TACCGACGTA
```

```
      +2 Y   A   A   Q   G   Y   K   V   L   V   L   N   P   S   V   A   A   T   L   G   F   G   A   Y   M   S   K
2001  ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA ACACTGGGCT TTGGTGCTTA CATGTCCAAG
      TACGTCGAGT CCCGATATTC CACGATCATG AGTTGGGGAG ACAACGACGT TGTGACCCGA AACCACGAAT GTACAGGTTC
```

```
      +2 A   H   G   I   D   P   N   I   R   T   G   V   R   T   I   T   T   G   S   P   I   T   Y   S   T   Y   G
2081  GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAC TGGCAGCCCC ATCACGTACT CCACCTACGG
      CGAGTACCCT AGCTAGGATT GTAGTCCTGG CCCCACTCTT GTTAATGGTG ACCGTCGGGG TAGTGCATGA GGTGGATGCC
```

```
      +2 K   F   L   A   D   G   G   C   S   G   G   A   Y   D   I   I   I   C   D   E   C   H   S   T   D   A
2161  CAAGTTCCTT GCCGACGGCG GGTGCTCGGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA
      GTTCAAGGAA CGGCTGCCGC CCACGAGCCC CCCGCGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCCTACGGT
```

```
      +2 T   S   I   L   G   I   G   T   V   L   D   Q   A   E   T   A   G   A   R   L   V   V   L   A   T   A   T
2241  CATCCATCTT GGGCATTGGC ACTGTCCTTG ACCAAGCAGA GACTGCGGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC
      GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGGTTCGTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG
```

```
      +2 P   P   G   S   V   T   V   P   H   P   N   I   E   E   V   A   L   S   T   T   G   E   I   P   F   Y   G
2321  CCTCCGGGCT CCGTCACTGT GCCCCATCCC AACATCGAGG AGGTTGCTCT GTCCACCACC GGAGAGATCC CTTTTTACGG
      GGAGGCCCGA GGCAGTGACA CGGGGTAGGG TTGTAGCTCC TCCAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC
```

```
      +2 K   A   I   P   L   E   V   I   K   G   G   R   H   L   I   F   C   H   S   K   K   K   C   D   E   L
2401  CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTTCT GTCATTCAAA GAAGAAGTGC GACGAACTCG
      GTTCCGATAG GGGGAGCTTC ATTAGTTCCC CCCCTCTGTA GAGTAGAAGA CAGTAAGTTT CTTCTTCACG CTGCTTGAGC
```

```
      +2 A   A   K   L   V   A   L   G   I   N   A   V   A   Y   Y   R   G   L   D   V   S   V   I   P   T   S   G
2481  CCGCAAAGCT GGTCGCATTG GGCATCAATG CCGTGGCCTA CTACCGCGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC
      GGCGTTTCGA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGCGCCA GAACTGCACA GGCAGTAGGG CTGGTCGCCG
```

```
      +2 D   V   V   V   V   A   T   D   A   L   M   T   G   Y   T   G   D   F   D   S   V   I   D   C   N   T   C
2561  GATGTTGTCG TCGTGGCAAC CGATGCCCTC ATGACCGGCT ATACCGGCGA CTTCGACTCG GTGATAGACT GCAATACGTG
      CTACAACAGC AGCACCGTTG GCTACGGGAG TACTGGCCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC
```

```
      +2   V  T  Q   T  V  D  F   S  L  D   P  T  F   T  I  E  T   I  T  L   P  Q  D   A  V  S
2641 . TGTCACCCAG ACAGTCGATT TCAGCCTTGA CCCTACCTTC ACCATTGAGA CAATCACGCT CCCCCAAGAT GCTGTCTCCC
       ACAGTGGGTC TGTCAGCTAA AGTCGGAACT GGGATGGAAG TGGTAACTCT GTTAGTGCGA GGGGGTTCTA CGACAGAGGG
```

```
      +2   R  T  Q  R   R  G  R   T  G  R  G   K  P  G   I  Y  R   F  V  A  P   G  E  R   P  S  G
2721   GCACTCAACG TCGGGGCAGG ACTGGCAGGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGGC
       CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AAACACCGTG GCCCCCTCGC GGGGAGGCCG
```

```
      +2   M  F  D  S   S  V  L   C  E  C   Y  D  A  G   C  A  W   Y  E  L   T  P  A  E   T  T  V
2801   ATGTTCGACT CGTCCGTCCT CTGTGAGTGC TATGACGCAG GCTGTGCTTG GTATGAGCTC ACGCCCGCCG AGACTACAGT
       TACAAGCTGA GCAGGCAGGA GACACTCACG ATACTGCGTC CGACACGAAC CATACTCGAG TGCGGGCGGC TCTGATGTCA
```

```
      +2   R  L  R   A  Y  M  N   T  P  G   L  P  V   C  Q  D  H   L  E  F   W  E  G   V  F  T
                                                                                            StuI
                                                                                            ~~
2881   TAGGCTACGA GCGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG
       ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGGTCCTGG TAGAACTTAA AACCCTCCCG CAGAAATGTC
```

```
      +2   G  L  T  H   I  D  A   H  F  L  S   Q  T  K   Q  S  G   E  N  L  P   Y  L  V   A  Y  Q
           StuI
           ~~~~
2961   GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAA
       CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTTT CGTCTCACCC CTCTTGGAAG GAATGGACCA TCGCATGGTT
```

```
      +2   A  T  V  C   A  R  A   Q  A  P   P  P  S  W   D  Q  M   W  K  C   L  I  R  L   K  P  T
3041   GCCACCGTGT GCGCTAGGGC TCAAGCCCCT CCCCCATCGT GGGACCAGAT GTGGAAGTGT TTGATTCGCC TCAAGCCCAC
       CGGTGGCACA CGCGATCCCG AGTTCGGGGA GGGGGTAGCA CCCTGGTCTA CACCTTCACA AACTAAGCGG AGTTCGGGTG
```

```
      +2   L  H  G   P  T  P  L   L  Y  R   L  G  A   V  Q  N  E   I  T  L   T  H  P   V  T  K
3121   CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGGCGCT GTTCAGAATG AAATCACCCT GACGCACCCA GTCACCAAAT
       GGAGGTACCC GGTTGTGGGG ACGATATGTC TGACCCGCGA CAAGTCTTAC TTTAGTGGGA CTGCGTGGGT CAGTGGTTTA
```

```
      +2   Y  I  M  T   C  M  S   A  D  L  E   V  V  T   S  T  W   V  L  V  G   G  V  L   A  A  L
3201   ACATCATGAC ATGCATGTCG GCCGACCTGG AGGTCGTCAC GAGCACCTGG GTGCTCGTTG GCGGCGTCCT GGCTGCTTTG
       TGTAGTACTG TACGTACAGC CGGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC CGCCGCAGGA CCGACGAAAC
```

```
      +2   A  A  Y  C   L  S  T   G  C  V   V  I  V  G   R  V  V   L  S  G   K  P  A  I   I  P  D
3281   GCCGCGTATT GCCTGTCAAC AGGCTGCGTG GTCATAGTGG GCAGGGTCGT CTTGTCCGGG AAGCCGGCAA TCATACCTGA
       CGGCGCATAA CGGACAGTTG TCCGACGCAC CAGTATCACC CGTCCCAGCA GAACAGGCCC TTCGGCCGTT AGTATGGACT
```

```
      +2   R  E  V   L  Y  R  E   F  D  E   M  E  E   C  S  Q  H   L  P  Y   I  E  Q   G  M  M
3361   CAGGGAAGTC CTCTACCGAG AGTTCGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC
       GTCCCTTCAG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTCG TGAATGGCAT GTAGCTCGTT CCCTACTACG
```

```
      +2  L  A  E  Q   F  K  Q   K  A  L  G   L  L  Q   T  A  S   R  Q  A  E   V  I  A   P  A  V
3441   TCGCCGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGCGTCC CGTCAGGCAG AGGTTATCGC CCCTGCTGTC
       AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG
```

```
      +2   Q  T  N  W   Q  K  L   E  T  F   W  A  K  H   M  W  N   F  I  S   G  I  Q  Y   L  A  G
3521   CAGACCAACT GGCAAAAACT CGAGACCTTC TGGGCGAAGC ATATGTGGAA CTTCATCAGT GGGATACAAT ACTTGGCGGG
       GTCTGGTTGA CCGTTTTTGA GCTCTGGAAG ACCCGCTTCG TATACACCTT GAAGTAGTCA CCCTATGTTA TGAACCGCCC
```

```
      +2   L  S  T   L  P  G  N   P  A  I   A  S  L   M  A  F  T   A  A  V   T  S  P   L  T  T
3601   CTTGTCAACG CTGCCTGGTA ACCCCGCCAT TGCTTCATTG ATGGCTTTTA CAGCTGCTGT CACCAGCCCA CTAACCACTA
       GAACAGTTGC GACGGACCAT TGGGGCGGTA ACGAAGTAAC TACCGAAAAT GTCGACGACA GTGGTCGGGT GATTGGTGAT
```

```
      +2   S  Q  T  L   L  F  N   I  L  G  G   W  V  A   A  Q  L   A  A  P  G   A  A  T   A  F  V
3681   GCCAAACCCT CCTCTTCAAC ATATTGGGGG GGTGGGTGGC TGCCCAGCTC GCCGCCCCCG GTGCCGCTAC TGCCTTTGTG
       CGGTTTGGGA GGAGAAGTTG TATAACCCCC CCACCCACCG ACGGGTCGAG CGGCGGGGGC CACGGCGATG ACGGAAACAC
```

FIGURE 3 - Page 4

```
       +2   G  A  G  L  A  G  A  A  I  G  S  V  G  L  G  K  V   L  I  D   I  L  A  G   Y  G  A
     3761   GGCGCTGGCT TAGCTGGCGC CGCCATCGGC AGTGTTGGAC TGGGGAAGGT CCTCATAGAC ATCCTTGCAG GGTATGGCGC
            CCGCGACCGA ATCGACCGCG GCGGTAGCCG TCACAACCTG ACCCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGCG

       +2   G  V  A  G  A  L  V  A  F  K  I  M  S  G  E  V  P  S  T  E   D  L  V  N  L  L
     3841   GGGCGTGGCG GGAGCTCTTG TGGCATTCAA GATCATGAGC GGTGAGGTCC CCTCCACGGA GGACCTGGTC AATCTACTGC
            CCCGCACCGC CCTCGAGAAC ACCGTAAGTT CTAGTACTCG CCACTCCAGG GGAGGTGCCT CCTGGACCAG TTAGATGACG

       +2 P  A  I  L  S  P  G  A  L  V  V  G  V  V  C  A  A  I  L  R  R   H  V  G  P  G  E
     3921   CCGCCATCCT CTCGCCCGGA GCCCTCGTAG TCGGCGTGGT CTGTGCAGCA ATACTGCGCC GGCACGTTGG CCCGGGCGAG
            GGCGGTAGGA GAGCGGGCCT CGGGAGCATC AGCCGCACCA GACACGTCGT TATGACGCGG CCGTGCAACC GGGCCCGCTC

       +2   G  A  V  Q  W  M  N  R  L  I  A  F  A  S  R  G  N  H  V  S  P  T  H  Y  V  P  E
     4001   GGGGCAGTGC AGTGGATGAA CCGGCTGATA GCCTTCGCCT CCCGGGGGAA CCATGTTTCC CCCACGCACT ACGTGCCGGA
            CCCCGTCACG TCACCTACTT GGCCGACTAT CGGAAGCGGA GGGCCCCCTT GGTACAAAGG GGGTGCGTGA TGCACGGCCT

       +2   S  D  A  A  A  R  V  T  A  I  L  S  S  L  T  V  T  Q  L  L  R  R  L  H  Q  W
     4081   GAGCGATGCA GCTGCCCGCG TCACTGCCAT ACTCAGCAGC CTCACTGTAA CCCAGCTCCT GAGGCGACTG CACCAGTGGA
            CTCGCTACGT CGACGGGCGC AGTGACGGTA TGAGTCGTCG GAGTGACATT GGGTCGAGGA CTCCGCTGAC GTGGTCACCT

       +2   I  S  S  E  C  T  T  P  C  S  G  S  W  L  R  D  I  W  D  W  I  C  E  V  L  S  D
     4161   TAAGCTCGGA GTGTACCACT CCATGCTCCG GTTCCTGGCT AAGGGACATC TGGGACTGGA TATGCGAGGT GTTGAGCGAC
            ATTCGAGCCT CACATGGTGA GGTACGAGGC CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACTCGCTG

       +2   F  K  T  W  L  K  A  K  L  M  P  Q  L  P  G  I  P  F  V  S  C  Q  R  G  Y  K  G
                                                       BamHI
                                                       ~~~~~~
     4241   TTTAAGACCT GGCTAAAAGC TAAGCTCATG CCACAGCTGC CTGGGATCCC CTTTGTGTCC TGCCAGCGCG GGTATAAGGG
            AAATTCTGGA CCGATTTTCG ATTCGAGTAC GGTGTCGACG GACCCTAGGG GAAACACAGG ACGGTCGCGC CCATATTCCC

       +2   V  W  R  G  D  G  I  M  H  T  R  C  H  C  G  A  E  I  T  G  H  V  K  N  G  T
     4321   GGTCTGGCGA GGGGACGGCA TCATGCACAC TCGCTGCCAC TGTGGAGCTG AGATCACTGG ACATGTCAAA AACGGGACGA
            CCAGACCGCT CCCCTGCCGT AGTACGTGTG AGCGACGGTG ACACCTCGAC TCTAGTGACC TGTACAGTTT TTGCCCTGCT

       +2 M  R  I  V  G  P  R  T  C  R  N  M  W  S  G  T  F  P  I  N  A  Y  T  T  G  P  C
     4401   TGAGGATCGT CGGTCCTAGG ACCTGCAGGA ACATGTGGAG TGGGACCTTC CCCATTAATG CCTACACCAC GGGCCCCTGT
            ACTCCTAGCA GCCAGGATCC TGGACGTCCT TGTACACCTC ACCCTGGAAG GGGTAATTAC GGATGTGGTG CCCGGGGACA

       +2   T  P  L  P  A  P  N  Y  T  F  A  L  W  R  V  S  A  E  E  Y  V  E  I  R  Q  V  G
     4481   ACCCCCCTTC CTGCGCCGAA CTACACGTTC GCGCTATGGA GGGTGTCTGC AGAGGAATAC GTGGAGATAA GGCAGGTGGG
            TGGGGGGAAG GACGCGGCTT GATGTGCAAG CGCGATACCT CCCACAGACG TCTCCTTATG CACCTCTATT CCGTCCACCC

       +2   D  F  H  Y  V  T  G  M  T  T  D  N  L  K  C  P  C  Q  V  P  S  P  E  F  F  T
     4561   GGACTTCCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCCGT GCCAGGTCCC ATCGCCCGAA TTTTTCACAG
            CCTGAAGGTG ATGCACTGCC CATACTGATG ACTGTTAGAA TTTACGGGCA CGGTCCAGGG TAGCGGGCTT AAAAAGTGTC

       +2   E  L  D  G  V  R  L  H  R  F  A  P  P  C  K  P  L  L  R  E  E  V  S  F  R  V  G
     4641   AATTGGACGG GGTGCGCCTA CATAGGTTTG CGCCCCCCTG CAAGCCCTTG CTGCGGGAGG AGGTATCATT CAGAGTAGGA
            TTAACCTGCC CCACGCGGAT GTATCCAAAC GCGGGGGGAC GTTCGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCCT

       +2   L  H  E  Y  P  V  G  S  Q  L  P  C  E  P  E  P  D  V  A  V  L  T  S  M  L  T  D
     4721   CTCCACGAAT ACCCGGTAGG GTCGCAATTA CCTTGCGAGC CCGAACCGGA CGTGGCCGTG TTGACGTCCA TGCTCACTGA
            GAGGTGCTTA TGGGCCATCC CAGCGTTAAT GGAACGCTCG GGCTTGGCCT GCACCGGCAC AACTGCAGGT ACGAGTGACT

       +2   P  S  H  I  T  A  E  A  A  G  R  R  L  A  R  G  S  P  P  S  V  A  S  S  S  A
     4801   TCCCTCCCAT ATAACAGCAG AGGCGGCCGG GCGAAGGTTG GCGAGGGGAT CACCCCCCTC TGTGGCCAGC TCCTCGGCTA
            AGGGAGGGTA TATTGTCGTC TCCGCCGGCC CGCTTCCAAC CGCTCCCCTA GTGGGGGGAG ACACCGGTCG AGGAGCCGAT
```

```
     +2 S  Q  L  S   A  P  S   L  K  A  T   C  T  A   N  H  D   S  P  D  A   E  L  I   E  A  N
   4881 GCCAGCTATC CGCTCCATCT CTCAAGGCAA CTTGCACCGC TAACCATGAC TCCCCTGATG CTGAGCTCAT AGAGGCCAAC
        CGGTCGATAG GCGAGGTAGA GAGTTCCGTT GAACGTGGCG ATTGGTACTG AGGGGACTAC GACTCGAGTA TCTCCGGTTG

     +2 L  L  W  R   Q  E  M   G  G  N   I  T  R  V   E  S  E   N  K  V   V  I  L   D  S  F  D
   4961 CTCCTATGGA GGCAGGAGAT GGGCGGCAAC ATCACCAGGG TTGAGTCAGA AAACAAAGTG GTGATTCTGG ACTCCTTCGA
        GAGGATACCT CCGTCCTCTA CCCGCCGTTG TAGTGGTCCC AACTCAGTCT TTTGTTTCAC CACTAAGACC TGAGGAAGCT

     +2 P  L  V   A  E  E  D   E  R  E   I  S  V   P  A  E  I   L  R  K   S  R  R   F  A  Q
   5041 TCCGCTTGTG GCGGAGGAGG ACGAGCGGGA GATCTCCGTA CCCGCAGAAA TCCTGCGGAA GTCTCGGAGA TTCGCCCAGG
        AGGCGAACAC CGCCTCCTCC TGCTCGCCCT CTAGAGGCAT GGGCGTCTTT AGGACGCCTT CAGAGCCTCT AAGCGGGTCC

     +2 A  L  P  V   W  A  R   P  D  Y  N   P  P  L   V  E  T   W  K  K  P   D  Y  E   P  P  V
   5121 CCCTGCCCGT TTGGGCGCGG CCGGACTATA ACCCCCCGCT AGTGGAGACG TGGAAAAAGC CGACTACGA ACCACCTGTG
        GGGACGGGCA AACCCGCGCC GGCCTGATAT TGGGGGGCGA TCACCTCTGC ACCTTTTTCG GCTGATGCT TGGTGGACAC

     +2 V  H  G  C   P  L  P   P  P  K   S  P  P  V   P  P  P   R  K  K   R  T  V  V   L  T  E
   5201 GTCCATGGCT GCCCGCTTCC ACCTCCAAAG TCCCCTCCTG TGCCTCCGCC TCGGAAGAAG CGGACGGTGG TCCTCACTGA
        CAGGTACCGA CGGGCGAAGG TGGAGGTTTC AGGGGAGGAC ACGGAGGCGG AGCCTTCTTC GCCTGCCACC AGGAGTGACT

     +2 S  T  L   S  T  A  L   A  E  L   A  T  R   S  F  G  S   S  S  T   S  G  I   T  G  D
   5281 ATCAACCCTA TCTACTGCCT TGGCCGAGCT CGCCACCAGA AGCTTTGGCA GCTCCTCAAC TTCCGGCATT ACGGGCGACA
        TAGTTGGGAT AGATGACGGA ACCGGCTCGA GCGGTGGTCT TCGAAACCGT CGAGGAGTTG AAGGCCGTAA TGCCCGCTGT

     +2 N  T  T  T   S  S  E   P  A  P  S   G  C  P   P  D  S   D  A  E  S   Y  S  S   M  P  P
   5361 ATACGACAAC ATCCTCTGAG CCCGCCCCTT CTGGCTGCCC CCCCGACTCC GACGCTGAGT CCTATTCCTC CATGCCCCCC
        TATGCTGTTG TAGGAGACTC GGGCGGGGAA GACCGACGGG GGGGCTGAGG CTGCGACTCA GGATAAGGAG GTACGGGGGG

     +2 L  E  G  E   P  G  D   P  D  L   S  D  G  S   W  S  T   V  S  S   E  A  N  A   E  D  V
                            BamHI
                            ~~~~~~~
   5441 CTGGAGGGGG AGCCTGGGGA TCCGGATCTT AGCGACGGGT CATGGTCAAC GGTCAGTAGT GAGGCCAACG CGGAGGATGT
        GACCTCCCCC TCGGACCCCT AGGCCTAGAA TCGCTGCCCA GTACCAGTTG CCAGTCATCA CTCCGGTTGC GCCTCCTACA

     +2 V  C  C   S  M  S  Y   S  W  T   G  A  L   V  T  P  C   A  A  E   E  Q  K   L  P  I
   5521 CGTGTGCTGC TCAATGTCTT ACTCTTGGAC AGGCGCACTC GTCACCCCGT GCGCCGCGGA AGAACAGAAA CTGCCCATCA
        GCACACGACG AGTTACAGAA TGAGAACCTG TCCGCGTGAG CAGTGGGGCA CGCGGCGCCT TCTTGTCTTT GACGGGTAGT

     +2 N  A  L  S   N  S  L   L  R  H  H   N  L  V   Y  S  T   T  S  R  S   A  C  Q   R  Q  K
   5601 ATGCACTAAG CAACTCGTTG CTACGTCACC ACAATTTGGT GTATTCCACC ACCTCACGCA GTGCTTGCCA AAGGCAGAAG
        TACGTGATTC GTTGAGCAAC GATGCAGTGG TGTTAAACCA CATAAGGTGG TGGAGTGCGT CACGAACGGT TTCCGTCTTC

     +2 K  V  T  F   D  R  L   Q  V  L   D  S  H  Y   Q  D  V   L  K  E   V  K  A  A   S  K
   5681 AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA
        TTTCAGTGTA AACTGTCTGA CGTTCAAGAC CTGTCGGTAA TGGTCCTGCA TGAGTTCCTC CAATTTCGTC GCCGCAGTTT

     +2 V  K  A   N  L  L  S   V  E  E   A  C  S   L  T  P  P   H  S  A   K  S  K   F  G  Y
   5761 AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGCAGC CTGACGCCCC CACACTCAGC CAAATCCAAG TTTGGTTATG
        TCACTTCCGA TTGAACGATA GGCATCTCCT TCGAACGTCG GACTGCGGGG GTGTGAGTCG GTTTAGGTTC AAACCAATAC

     +2 G  A  K  D   V  R  C   H  A  R  K   A  V  T   H  I  N   S  V  W  K   D  L  L   E  D  N
   5841 GGGCAAAAGA CGTCCGTTGC CATGCCAGAA AGGCCGTAAC CCACATCAAC TCCGTGTGGA AAGACCTTCT GGAAGACAAT
        CCCGTTTTCT GCAGGCAACG GTACGGTCTT TCCGGCATTG GGTGTAGTTG AGGCACACCT TTCTGGAAGA CCTTCTGTTA

     +2 V  T  P  I   D  T  T   I  M  A   K  N  E  V   F  C  V   Q  P  E   K  G  G  R   K  P  A
   5921 GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGCGT TCAGCCTGAG AAGGGGGGTC GTAAGCCAGC
        CATTGTGGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGACGCA AGTCGGACTC TTCCCCCCAG CATTCGGTCG
```

```
       +2   R   L   I   V   F   P   D   L   G   V   R   V   C   E   K   M   A   L   Y   D   V   V   T   K   L   P
     6001  TCGTCTCATC GTGTTCCCCG ATCTGGGCGT GCGCGTGTGC GAAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCCT
           AGCAGAGTAG CACAAGGGGC TAGACCCGCA CGCGCACACG CTTTTCTACC GAAACATGCT GCACCAATGT TTCGAGGGGA


       +2   L   A   V   M   G   S   S   Y   G   F   Q   Y   S   P   G   Q   R   V   E   F   L   V   Q   A   W   K   S
                                                                            EcoRI
                                                                           ~~~~~~
     6081  TGGCCGTGAT GGGAAGCTCC TACGGATTCC AATACTCACC AGGACAGCGG GTTGAATTCC TCGTGCAAGC GTGGAAGTCC
           ACCGGCACTA CCCTTCGAGG ATGCCTAAGG TTATGAGTGG TCCTGTCGCC CAACTTAAGG AGCACGTTCG CACCTTCAGG


       +2   K   K   T   P   M   G   F   S   Y   D   T   R   C   F   D   S   T   V   T   E   S   D   I   R   T   E   E
     6161  AAGAAAACCC CAATGGGGTT CTCGTATGAT ACCCGCTGCT TTGACTCCAC AGTCACTGAG AGCGACATCC GTACGGAGGA
           TTCTTTTGGG GTTACCCCAA GAGCATACTA TGGGCGACGA AACTGAGGTG TCAGTGACTC TCGCTGTAGG CATGCCTCCT


       +2   A   I   Y   Q   C   C   D   L   D   P   Q   A   R   V   A   I   K   S   L   T   E   R   L   Y   V   G
     6241  GGCAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG
           CCGTTAGATG GTTACAACAC TGGAGCTGGG GGTTCGGGCG CACCGGTAGT TCAGGGAGTG GCTCTCCGAA ATACAACCCC


       +2   G   P   L   T   N   S   R   G   E   N   C   G   Y   R   R   C   R   A   S   G   V   L   T   T   S   C   G
     6321  GCCCTCTTAC CAATTCAAGG GGGGAGAACT GCGGCTATCG CAGGTGCCGC GCGAGCGGCG TACTGACAAC TAGCTGTGGT
           CGGGAGAATG GTTAAGTTCC CCCCTCTTGA CGCCGATAGC GTCCACGGCG CGCTCGCCGC ATGACTGTTG ATCGACACCA


       +2   N   T   L   T   C   Y   I   K   A   R   A   A   C   R   A   A   G   L   Q   D   C   T   M   L   V   C   G
     6401  AACACCCTCA CTTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCGCAGG GCTCCAGGAC TGCACCATGC TCGTGTGTGG
           TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCGTCC CGAGGTCCTG ACGTGGTACG AGCACACACC


       +2   D   D   L   V   V   I   C   E   S   A   G   V   Q   E   D   A   A   S   L   R   A   F   T   E   A   M
     6481  CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGGGTCCAG GAGGACGCGG CGAGCCTGAG AGCCTTCACG GAGGCTATGA
           GCTGCTGAAT CAGCAATAGA CACTTTCGCG CCCCCAGGTC CTCCTGCGCC GCTCGGACTC TCGGAAGTGC CTCCGATACT


       +2   T   R   Y   S   A   P   P   G   D   P   P   Q   P   E   Y   D   L   E   L   I   T   S   C   S   S   N   V
     6561  CCAGGTACTC CGCCCCCCCT GGGGACCCCC CACAACCAGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGTG
           GGTCCATGAG GCGGGGGGGA CCCCTGGGGG GTGTTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC


       +2   S   V   A   H   D   G   A   G   K   R   V   Y   Y   L   T   R   D   P   T   T   P   L   A   R   A   A   W
     6641  TCAGTCGCCC ACGACGGCGC TGGAAAGAGG GTCTACTACC TCACCCGTGA CCCTACAACC CCCCTCGCGA GAGCTGCGTG
           AGTCAGCGGG TGCTGCCGCG ACCTTTCTCC CAGATGATGG AGTGGGCACT GGGATGTTGG GGGGAGCGCT CTCGACGCAC


       +2   E   T   A   R   H   T   P   V   N   S   W   L   G   N   I   I   M   F   A   P   T   L   W   A   R   M
     6721  GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTTGCCCC CACACTGTGG GCGAGGATGA
           CCTCTGTCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAAACGGGG GTGTGACACC CGCTCCTACT


       +2   I   L   M   T   H   F   F   S   V   L   I   A   R   D   Q   L   E   Q   A   L   D   C   E   I   Y   G   A
     6801  TACTGATGAC CCATTTCTTT AGCGTCCTTA TAGCCAGGGA CCAGCTTGAA CAGGCCCTCG ATTGCGAGAT CTACGGGGCC
           ATGACTACTG GGTAAAGAAA TCGCAGGAAT ATCGGTCCCT GGTCGAACTT GTCCGGGAGC TAACGCTCTA GATGCCCCGG


       +2   C   Y   S   I   E   P   L   D   L   P   P   I   I   Q   R   L   H   G   L   S   A   F   S   L   H   S   Y
     6881  TGCTACTCCA TAGAACCACT GGATCTACCT CCAATCATTC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA
           ACGATGAGGT ATCTTGGTGA CCTAGATGGA GGTTAGTAAG TTTCTGAGGT ACCGGAGTCG CGTAAAAGTG AGGTGTCAAT


       +2   S   P   G   E   I   N   R   V   A   A   C   L   R   K   L   G   V   P   P   L   R   A   W   R   H   R
     6961  CTCTCCAGGT GAAATCAATA GGGTGGCCGC ATGCCTCAGA AAACTTGGGG TACCGCCCTT GCGAGCTTGG AGACACCGGG
           GAGAGGTCCA CTTTAGTTAT CCCACCGGCG TACGGAGTCT TTTGAACCCC ATGGCGGGAA CGCTCGAACC TCTGTGGCCC


       +2   A   R   S   V   R   A   R   L   L   A   R   G   G   R   A   A   I   C   G   K   Y   L   F   N   W   A   V
     7041  CCCGGAGCGT CCGCGCTAGG CTTCTGGCCA GAGGAGGCAG GGCTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA
           GGGCCTCGCA GGCGCGATCC GAAGACCGGT CTCCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCCGTCAT
```

```
     +2  R   T  K  L   K  L  T   P  I  A   A  A  G  Q   L  D  L   S  G  W   F  T  A  G   Y  S  G
7121  AGAACAAAGC TCAAACTCAC TCCAATAGCG GCCGCTGGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG
      TCTTGTTTCG AGTTTGAGTG AGGTTATCGC CGGCGACCGG TCGACCTGAA CAGGCCGACC AAGTGCCGAC CGATGTCGCC
```

```
     +2  G   D  I  Y  H  S  V   S  H  A   R  P  R   W  I  W  F   C  L  L   L  L  A   A  G  V
7201  GGGAGACATT TATCACAGCG TGTCTCATGC CCGGCCCCGC TGGATCTGGT TTTGCCTACT CCTGCTTGCT GCAGGGGTAG
      CCCTCTGTAA ATAGTGTCGC ACAGAGTACG GGCCGGGGCG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCCCATC
```

```
     +2 G   I  Y  L   L  P  N   R
7281  GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGGTAAAC ACTCCGGCCT AAAAAAAAAA AAAAATCTAG AAAGGCGCGC
      CGTAGATGGA GGAGGGGTTG GCTACTTCCA ACCCCATTTG TGAGGCCGGA TTTTTTTTTT TTTTTAGATC TTTCCGCGCG
```

```
                      .BamHI     MluI
                      ~~~~~~     ~~~~~~
7361  CAAGATATCA AGGATCCACT ACGCGTTAGA GCTCGCTGAT CAGCCTCGAC TGTGCCTTCT AGTTGCCAGC CATCTGTTGT
      GTTCTATAGT TCCTAGGTGA TGCGCAATCT CGAGCGACTA GTCGGAGCTG ACACGGAAGA TCAACGGTCG GTAGACAACA
```

```
7441  TTGCCCCTCC CCCGTGCCTT CCTTGACCCT GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAAATGAG GAAATTGCAT
      AACGGGGAGG GGGCACGGAA GGAACTGGGA CCTTCCACGG TGAGGGTGAC AGGAAAGGAT TATTTTACTC CTTTAACGTA
```

```
7521  CGCATTGTCT GAGTAGGTGT CATTCTATTC TGGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT
      GCGTAACAGA CTCATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTCGTTCC CCCTCCTAAC CCTTCTGTTA
```

```
7601  AGCAGGCATG CTGGGGAGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT
      TCGTCCGTAC GACCCCTCGA GAAGGCGAAG GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCCGACGCC GCTCGCCATA
```

```
7681  CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG
      GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCTATTG CGTCCTTTCT TGTACACTCG TTTTCCGGTC
```

```
7761  CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA
      GTTTTCCGGT CCTTGGCATT TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCGGGG GGACTGCTCG TAGTGTTTTT
```

```
7841  TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC
      AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCCTGAT ATTTCTATGG TCCGCAAAGG GGGACCTTCG AGGGAGCACG
```

```
7921  GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCAATGC
      CGAGAGGACA AGGCTGGGAC GGCGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCGA AAGAGTTACG
```

```
8001  TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA
      AGTGCGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GGTTCGACCC GACACACGTG CTTGGGGGGC AAGTCGGGCT
```

```
8081  CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG
      GGCGACGCGG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CGGTGACCGT CGTCGGTGAC
```

```
8161  GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACTACGG CTACACTAGA
      CATTGTCCTA ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT
```

```
8241  AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAACA
      TCCTGTCATA AACCATAGAC GCGAGACGAC TTCGGTCAAT GGAAGCCTTT TTCTCAACCA TCGAGAACTA GGCCGTTTGT
```

```
8321  AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT
      TTGGTGGCGA CCATCGCCAC CAAAAAAACA AACGTTCGTC GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTTCTAGGAA
```

```
8401  TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG
      ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT CCCTAAAACC AGTACTCTAA TAGTTTTTCC
```

```
8481  ATCTTCACCT AGATCCTTTT AAAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG
      .TAGAAGTGGA TCTAGGAAAA TTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA CCAGACTGTC

8561  TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTC GTTCATCCAT AGTTGCCTGA CTCCCCGTCG
      AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCAGC

8641  TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT
      ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACCGGG GTCACGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA

8721  CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA
      GGTCTAAATA GTCGTTATTT GGTCGGTCGG CCTTCCCGGC TCGCGTCTTC ACCAGGACGT TGAAATAGGC GGAGGTAGGT

8801  GTCTATTAAT TGTTGCCGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG
      CAGATAATTA ACAACGGCCC TTCGATCTCA TTCATCAAGC GGTCAATTAT CAAACGCGTT GCAACAACGG TAACGATGTC

8881  GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC
      CGTAGCACCA CAGTGCGAGC AGCAAACCAT ACCGAAGTAA GTCGAGGCCA AGGGTTGCTA GTTCCGCTCA ATGTACTAGG

8961  CCCATGTTGT GCAAAAAAGC GGTTAGCTCC TTCGGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTTATCACT
      GGGTACAACA CGTTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCATTC AACCGGCGTC ACAATAGTGA

9041  CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA
      GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCATT CTACGAAAAG ACACTGACCA CTCATGAGTT

9121  CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT
      GGTTCAGTAA GACTCTTATC ACATACGCCG CTGGCTCAAC GAGAACGGGC CGCAGTTATG CCCTATTATG GCGCGGTGTA

9201  AGCAGAACTT TAAAAGTGCT CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTTGAGATC
      TCGTCTTGAA ATTTTCACGA GTAGTAACCT TTTGCAAGAA GCCCCGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG

9281  CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA
      GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAAT GAAAGTGGTC GCAAAGACCC ACTCGTTTTT

9361  CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT
      GTCCTTCCGT TTTACGGCGT TTTTTCCCTT ATTCCCGCTG TGCCTTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA

9441  TATTGAAGCA TTTATCAGGG TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC AAATAGGGGT
      ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT TACATAAATC TTTTTATTTG TTTATCCCCA

9521  TCCGCGCACA TTTCCCCGAA AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAGGC
      AGGCGCGTGT AAAGGGGCTT TTCACGGTGG ACTGCAGATT CTTTGGTAAT AATAGTACTG TAATTGGATA TTTTTATCCG

9601  GTATCACGAG GCCCTTTCGT C
      CATAGTGCTC CGGGAAAGCA G
```

FIGURE 4

Stu I (211)

*Eco*R I (1983)

Mlu I (7382)
*Bam*H I (7373)

pCMV-delNS35
9621 bp

Stu I (2962)

*Eco*R I (6135)

*Bam*H I (5459)

*Bam*H I (4285)

delNS35 ORF
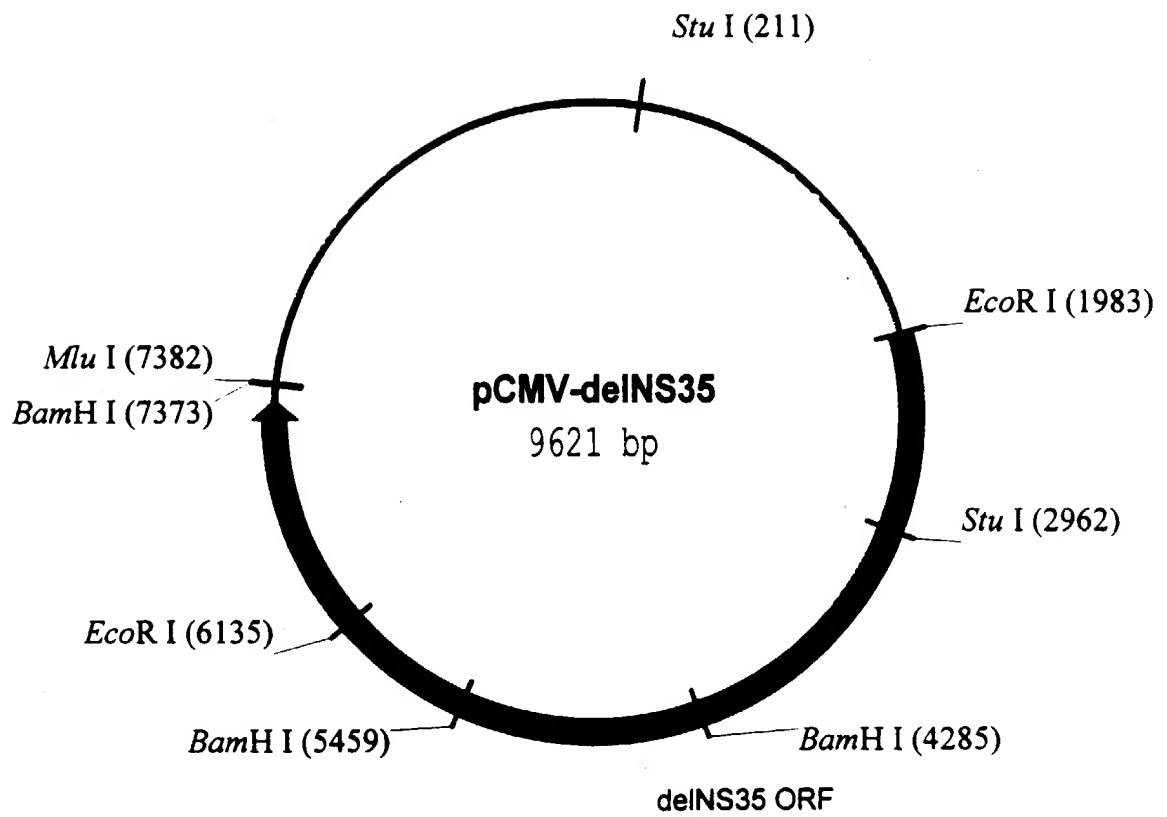
```
   1   TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT
       AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTCGCCTA

  81   GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA
       CGGCCCTCGT CTGTTCGGGC AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC GCCGTAGTCT

                                                                        StuI
                                                                        ~~~~~~~
 161   GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG
       CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC

 241   AATAGCTCAG AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGGCGGA
       TTATCGAGTC TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT ACCCCGCCTC TTACCCGCCT

 321   ACTGGGCGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
       TGACCCGCCC CTCCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

 401   CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT
       GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA

 481   AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCCATT
       TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA

 561   GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT
       CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

 641   AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG TCAATGACGG TAAATGGCCC
       TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC AGTTACTGCC ATTTACCGGG

 721   GCCTGGCATT ATGCCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
       CGGACCGTAA TACGGGTCAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

 801   CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA
       GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT

 881   TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC CCCGTTGACG
       AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

 961   CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG
       GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1041   CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG GGAACGGTGC ATTGGAACGC
       GGTAGGTGCG ACAAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCCGGC CCTTGCCACG TAACCTTGCG

1121   GGATTCCCCG TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
       CCTAAGGGGC ACGGTTCTCA CTGCATTCAT GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201   CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA
       GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

1281   TTGACCATTA TTGACCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACTAT
       AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

1361   CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT
       GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA
```

```
1441  TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT
      ATAAATGTTT AAGTGTATAT GTTGTTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT ATCGCACCCT AGAGGCTGTA


1521  CTCGGGTACG TGTTCCGGAC ATGGGCTCTT CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
      GAGCCCATGC ACAAGGCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT


1601  GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC
      CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTTACGG GTGGTGGTGG


1681  AGTGTGCCGC ACAAGGCCGT GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT
      TCACACGGCG TGTTCCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA CCTGCGTCTA


1761  GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT
      CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA


1841  TGCGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT
      ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA


      +2                                                                          M  A  A
                                                                    EcoRI
                                                                    ~~~~~~
1921  GACAGACTAA CAGACTGTTC CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCTGCAT
      CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGACGTA


      +2 Y  A  A  Q   G  Y  K   V  L  V  L   N  P  S   V  A  A   T  L  G  F   G  A  Y   M  S  K
2001  ATGCAGCTCA GGGCTATAAG GTGCTAGTAC TCAACCCCTC TGTTGCTGCA ACACTGGGCT TTGGTGCTTA CATGTCCAAG
      TACGTCGAGT CCCGATATTC CACGATCATG AGTTGGGGAG ACAACGACGT TGTGACCCGA AACCACGAAT GTACAGGTTC


      +2 A  H  G  I   D  P  N   I  R  T   G  V  R  T   I  T  T   G  S  P   I  T  Y  S   T  Y  G
2081  GCTCATGGGA TCGATCCTAA CATCAGGACC GGGGTGAGAA CAATTACCAC TGGCAGCCCC ATCACGTACT CCACCTACGG
      CGAGTACCCT AGCTAGGATT GTAGTCCTGG CCCCACTCTT GTTAATGGTG ACCGTCGGGG TAGTGCATGA GGTGGATGCC


      +2 K  F  L   A  D  G  G   C  S  G   G  A  Y   D  I  I  I   C  D  E   C  H  S   T  D  A
2161  CAAGTTCCTT GCCGACGGCG GGTGCTCGGG GGGCGCTTAT GACATAATAA TTTGTGACGA GTGCCACTCC ACGGATGCCA
      GTTCAAGGAA CGGCTGCCGC CCACGAGCCC CCCGCGAATA CTGTATTATT AAACACTGCT CACGGTGAGG TGCCTACGGT


      +2 T  S  I  L   G  I  G   T  V  L  D   Q  A  E   T  A  G   A  R  L  V   V  L  A   T  A  T
2241  CATCCATCTT GGGCATTGGC ACTGTCCTTG ACCAAGCAGA GACTGCGGGG GCGAGACTGG TTGTGCTCGC CACCGCCACC
      GTAGGTAGAA CCCGTAACCG TGACAGGAAC TGGTTCGTCT CTGACGCCCC CGCTCTGACC AACACGAGCG GTGGCGGTGG


      +2 P  P  G  S   V  T  V   P  H  P   N  I  E  E   V  A  L   S  T  T   G  E  I  P   F  Y  G
2321  CCTCCGGGCT CCGTCACTGT GCCCCATCCC AACATCGAGG AGGTTGCTCT GTCCACCACC GGAGAGATCC CTTTTTACGG
      GGAGGCCCGA GGCAGTGACA CGGGGTAGGG TTGTAGCTCC TCCAACGAGA CAGGTGGTGG CCTCTCTAGG GAAAAATGCC


      +2  K  A  I   P  L  E  V   I  K  G   G  R  H   L  I  F   C  H  S   K  K  C   D  E  L
2401  CAAGGCTATC CCCCTCGAAG TAATCAAGGG GGGGAGACAT CTCATCTTCT GTCATTCAAA GAAGAAGTGC GACGAACTCG
      GTTCCGATAG GGGGAGCTTC ATTAGTTCCC CCCCTCTGTA GAGTAGAAGA CAGTAAGTTT CTTCTTCACG CTGCTTGAGC


      +2 A  A  K  L   V  A  L   G  I  N  A   V  A  Y   Y  R  G   L  D  V  S   V  I  P   T  S  G
2481  CCGCAAAGCT GGTCGCATTG GGCATCAATG CCGTGGCCTA CTACCGCGGT CTTGACGTGT CCGTCATCCC GACCAGCGGC
      GGCGTTTCGA CCAGCGTAAC CCGTAGTTAC GGCACCGGAT GATGGCGCCA GAACTGCACA GGCAGTAGGG CTGGTCGCCG


      +2 D  V  V   V  A  T   D  A  L   M  T  G  Y   T  G  D   F  D  S   V  I  D  C   N  T  C
2561  GATGTTGTCG TCGTGGCAAC CGATGCCCTC ATGACCGGCT ATACCGGCGA CTTCGACTCG GTGATAGACT GCAATACGTG
      CTACAACAGC AGCACCGTTG GCTACGGGAG TACTGGCCGA TATGGCCGCT GAAGCTGAGC CACTATCTGA CGTTATGCAC
```

```
      +2     V  T  Q   T  V  D  F   S  L  D   P  T  F   T  I  E  T   I  T  L   P  Q  D   A  V  S
2641   TGTCACCCAG ACAGTCGATT TCAGCCTTGA CCCTACCTTC ACCATTGAGA CAATCACGCT CCCCCAAGAT GCTGTCTCCC
       ACAGTGGGTC TGTCAGCTAA AGTCGGAACT GGGATGGAAG TGGTAACTCT GTTAGTGCGA GGGGGTTCTA CGACAGAGGG
```

```
      +2     R  T  Q  R   R  G  R   T  G  R  G   K  P  G   I  Y  R   F  V  A  P   G  E  R   P  S  G
2721   GCACTCAACG TCGGGGCAGG ACTGGCAGGG GGAAGCCAGG CATCTACAGA TTTGTGGCAC CGGGGGAGCG CCCCTCCGGC
       CGTGAGTTGC AGCCCCGTCC TGACCGTCCC CCTTCGGTCC GTAGATGTCT AAACACCGTG GCCCCCTCGC GGGGAGGCCG
```

```
      +2     M  F  D  S   S  V  L   C  E  C   Y  D  A  G   C  A  W   Y  E  L   T  P  A  E   T  T  V
2801   ATGTTCGACT CGTCCGTCCT CTGTGAGTGC TATGACGCAG GCTGTGCTTG GTATGAGCTC ACGCCCGCCG AGACTACAGT
       TACAAGCTGA GCAGGCAGGA GACACTCACG ATACTGCGTC CGACACGAAC CATACTCGAG TGCGGGCGGC TCTGATGTCA
```

```
      +2     R  L  R   A  Y  M  N   T  P  G   L  P  V   C  Q  D  H   L  E  F   W  E  G   V  F  T
                                                                                          StuI
2881   TAGGCTACGA GCGTACATGA ACACCCCGGG GCTTCCCGTG TGCCAGGACC ATCTTGAATT TTGGGAGGGC GTCTTTACAG
       ATCCGATGCT CGCATGTACT TGTGGGGCCC CGAAGGGCAC ACGGTCCTGG TAGAACTTAA AACCCTCCCG CAGAAATGTC
```

```
      +2   G  L  T  H   I  D  A   H  F  L  S   Q  T  K   Q  S  G   E  N  L  P   Y  L  V   A  Y  Q
           StuI
           ~~~~
2961   GCCTCACTCA TATAGATGCC CACTTTCTAT CCCAGACAAA GCAGAGTGGG GAGAACCTTC CTTACCTGGT AGCGTACCAA
       CGGAGTGAGT ATATCTACGG GTGAAAGATA GGGTCTGTTT CGTCTCACCC CTCTTGGAAG GAATGGACCA TCGCATGGTT
```

```
      +2   A  T  V  C   A  R  A   Q  A  P   P  P  S  W   D  Q  M   W  K  C   L  I  R  L   K  P  T
3041   GCCACCGTGT GCGCTAGGGC TCAAGCCCCT CCCCCATCGT GGGACCAGAT GTGGAAGTGT TTGATTCGCC TCAAGCCCAC
       CGGTGGCACA CGCGATCCCG AGTTCGGGGA GGGGGTAGCA CCCTGGTCTA CACCTTCACA AACTAAGCGG AGTTCGGGTG
```

```
      +2   L  H  G   P  T  P  L   L  Y  R   L  G  A   V  Q  N  E   I  T  L   T  H  P   V  T  K
3121   CCTCCATGGG CCAACACCCC TGCTATACAG ACTGGGCGCT GTTCAGAATG AAATCACCCT GACGCACCCA GTCACCAAAT
       GGAGGTACCC GGTTGTGGGG ACGATATGTC TGACCCGCGA CAAGTCTTAC TTTAGTGGGA CTGCGTGGGT CAGTGGTTTA
```

```
      +2 Y  I  M  T   C  M  S   A  D  L  E   V  V  T   S  T  W   V  L  V  G   G  V  L   A  A  L
3201   ACATCATGAC ATGCATGTCG GCCGACCTGG AGGTCGTCAC GAGCACCTGG GTGCTCGTTG GCGGCGTCCT GGCTGCTTTG
       TGTAGTACTG TACGTACAGC CGGCTGGACC TCCAGCAGTG CTCGTGGACC CACGAGCAAC CGCCGCAGGA CCGACGAAAC
```

```
      +2   A  A  Y  C   L  S  T   G  C  V   V  I  V  G   R  V  V   L  S  G   K  P  A  I   I  P  D
3281   GCCGCGTATT GCCTGTCAAC AGGCTGCGTG GTCATAGTGG GCAGGGTCGT CTTGTCCGGG AAGCCGGCAA TCATACCTGA
       CGGCGCATAA CGGACAGTTG TCCGACGCAC CAGTATCACC CGTCCCAGCA GAACAGGCCC TTCGGCCGTT AGTATGGACT
```

```
      +2   R  E  V   L  Y  R  E   F  D  E   M  E  E   C  S  Q  H   L  P  Y   I  E  Q   G  M  M
3361   CAGGGAAGTC CTCTACCGAG AGTTCGATGA GATGGAAGAG TGCTCTCAGC ACTTACCGTA CATCGAGCAA GGGATGATGC
       GTCCCTTCAG GAGATGGCTC TCAAGCTACT CTACCTTCTC ACGAGAGTCG TGAATGGCAT GTAGCTCGTT CCCTACTACG
```

```
      +2 L  A  E  Q   F  K  Q   K  A  L  G   L  L  Q   T  A  S   R  Q  A  E   V  I  A   P  A  V
3441   TCGCCGAGCA GTTCAAGCAG AAGGCCCTCG GCCTCCTGCA GACCGCGTCC CGTCAGGCAG AGGTTATCGC CCCTGCTGTC
       AGCGGCTCGT CAAGTTCGTC TTCCGGGAGC CGGAGGACGT CTGGCGCAGG GCAGTCCGTC TCCAATAGCG GGGACGACAG
```

```
      +2   Q  T  N  W   Q  K  L   E  T  F   W  A  K  H   M  W  N   F  I  S   G  I  Q  Y   L  A  G
3521   CAGACCAACT GGCAAAAACT CGAGACCTTC TGGGCGAAGC ATATGTGGAA CTTCATCAGT GGGATACAAT ACTTGGCGGG
       GTCTGGTTGA CCGTTTTTGA GCTCTGGAAG ACCCGCTTCG TATACACCTT GAAGTAGTCA CCCTATGTTA TGAACCGCCC
```

```
      +2   L  S  T   L  P  G  N   P  A  I   A  S  L   M  A  F  T   A  A  V   T  S  P   L  T  T
3601   CTTGTCAACG CTGCCTGGTA ACCCCGCCAT TGCTTCATTG ATGGCTTTTA CAGCTGCTGT CACCAGCCCA CTAACCACTA
       GAACAGTTGC GACGGACCAT TGGGGCGGTA ACGAAGTAAC TACCGAAAAT GTCGACGACA GTGGTCGGGT GATTGGTGAT
```

```
      +2 S  Q  T  L   L  F  N   I  L  G  G   W  V  A   A  Q  L   A  A  P  G   A  A  T   A  F  V
3681   GCCAAACCCT CCTCTTCAAC ATATTGGGGG GGTGGGTGGC TGCCCAGCTC GCCGCCCCCG GTGCCGCTAC TGCCTTTGTG
       CGGTTTGGGA GGAGAAGTTG TATAACCCCC CCACCCACCG ACGGGTCGAG CGGCGGGGGC CACGGCGATG ACGGAAACAC
```

```
     +2  G  A  G  L  A  G  A  A  I  G   S  V  G  L  G  K  V   L  I  D  I  L  A  G   Y  G  A
3761 . GGCGCTGGCT TAGCTGGCGC CGCCATCGGC AGTGTTGGAC TGGGGAAGGT CCTCATAGAC ATCCTTGCAG GGTATGGCGC
       CCGCGACCGA ATCGACCGCG GCGGTAGCCG TCACAACCTG ACCCCTTCCA GGAGTATCTG TAGGAACGTC CCATACCGCG
```

```
     +2   G  V  A  G  A  L  V  A  F  K  I  M  S  G  E  V  P  S  T  E  D  L  V  N  L  L
3841 GGGCGTGGCG GGAGCTCTTG TGGCATTCAA GATCATGAGC GGTGAGGTCC CCTCCACGGA GGACCTGGTC AATCTACTGC
     CCCGCACCGC CCTCGAGAAC ACCGTAAGTT CTAGTACTCG CCACTCCAGG GGAGGTGCCT CCTGGACCAG TTAGATGACG
```

```
     +2 P  A  I  L  S  P  G  A  L  V  V  G  V  V  C  A  A  I  L  R  R  H  V  G  P  G  E
3921 CCGCCATCCT CTCGCCCGGA GCCCTCGTAG TCGGCGTGGT CTGTGCAGCA ATACTGCGCC GGCACGTTGG CCCGGGCGAG
     GGCGGTAGGA GAGCGGGCCT CGGGAGCATC AGCCGCACCA GACACGTCGT TATGACGCGG CCGTGCAACC GGGCCCGCTC
```

```
     +2 G  A  V  Q  W  M  N  R  L  I  A  F  A  S  R  G  N  H  V  S  P  T  H  Y  V  P  E
4001 GGGGCAGTGC AGTGGATGAA CCGGCTGATA GCCTTCGCCT CCCGGGGGAA CCATGTTTCC CCCACGCACT ACGTGCCGGA
     CCCCGTCACG TCACCTACTT GGCCGACTAT CGGAAGCGGA GGGCCCCCTT GGTACAAAGG GGGTGCGTGA TGCACGGCCT
```

```
     +2  S  D  A  A  A  R  V  T  A  I  L  S  S  L  T  V  T  Q  L  L  R  R  L  H  Q  W
4081 GAGCGATGCA GCTGCCCGCG TCACTGCCAT ACTCAGCAGC CTCACTGTAA CCCAGCTCCT GAGGCGACTG CACCAGTGGA
     CTCGCTACGT CGACGGGCGC AGTGACGGTA TGAGTCGTCG GAGTGACATT GGGTCGAGGA CTCCGCTGAC GTGGTCACCT
```

```
     +2 I  S  S  E  C  T  T  P  C  S  G  S  W  L  R  D  I  W  D  W  I  C  E  V  L  S  D
4161 TAAGCTCGGA GTGTACCACT CCATGCTCCG GTTCCTGGCT AAGGGACATC TGGGACTGGA TATGCGAGGT GTTGAGCGAC
     ATTCGAGCCT CACATGGTGA GGTACGAGGC CAAGGACCGA TTCCCTGTAG ACCCTGACCT ATACGCTCCA CAACTCGCTG
```

```
     +2  F  K  T  W  L  K  A  K  L  M  P  Q  L  P  G  I  P  F  V  S  C  Q  R  G  Y  K  G
                                                      BamHI
                                                     ~~~~~~
4241 TTTAAGACCT GGCTAAAAGC TAAGCTCATG CCACAGCTGC CTGGGATCCC CTTTGTGTCC TGCCAGCGCG GGTATAAGGG
     AAATTCTGGA CCGATTTTCG ATTCGAGTAC GGTGTCGACG GACCCTAGGG GAAACACAGG ACGGTCGCGC CCATATTCCC
```

```
     +2  V  W  R  G  D  G  I  M  H  T  R  C  H  C  G  A  E  I  T  G  H  V  K  N  G  T
4321 GGTCTGGCGA GGGGACGGCA TCATGCACAC TCGCTGCCAC TGTGGAGCTG AGATCACTGG ACATGTCAAA AACGGGACGA
     CCAGACCGCT CCCCTGCCGT AGTACGTGTG AGCGACGGTG ACACCTCGAC TCTAGTGACC TGTACAGTTT TTGCCCTGCT
```

```
     +2 M  R  I  V  G  P  R  T  C  R  N  M  W  S  G  T  F  P  I  N  A  Y  T  T  G  P  C
4401 TGAGGATCGT CGGTCCTAGG ACCTGCAGGA ACATGTGGAG TGGGACCTTC CCCATTAATG CCTACACCAC GGGCCCCTGT
     ACTCCTAGCA GCCAGGATCC TGGACGTCCT TGTACACCTC ACCCTGGAAG GGGTAATTAC GGATGTGGTG CCCGGGGACA
```

```
     +2 T  P  L  P  A  P  N  Y  T  F  A  L  W  R  V  S  A  E  E  Y  V  E  I  R  Q  V  G
4481 ACCCCCCTTC CTGCGCCGAA CTACACGTTC GCGCTATGGA GGGTGTCTGC AGAGGAATAC GTGGAGATAA GGCAGGTGGG
     TGGGGGGAAG GACGCGGCTT GATGTGCAAG CGCGATACCT CCCACAGACG TCTCCTTATG CACCTCTATT CCGTCCACCC
```

```
     +2   D  F  H  Y  V  T  G  M  T  T  D  N  L  K  C  P  C  Q  V  P  S  E  F  F  T
4561 GGACTTCCAC TACGTGACGG GTATGACTAC TGACAATCTT AAATGCCCGT GCCAGGTCCC ATCGCCCGAA TTTTTCACAG
     CCTGAAGGTG ATGCACTGCC CATACTGATG ACTGTTAGAA TTTACGGGCA CGGTCCAGGG TAGCGGGCTT AAAAAGTGTC
```

```
     +2 E  L  D  G  V  R  L  H  R  F  A  P  P  C  K  P  L  L  R  E  E  V  S  F  R  V  G
4641 AATTGGACGG GGTGCGCCTA CATAGGTTTG CGCCCCCCTG CAAGCCCTTG CTGCGGGAGG AGGTATCATT CAGAGTAGGA
     TTAACCTGCC CCACGCGGAT GTATCCAAAC GCGGGGGGAC GTTCGGGAAC GACGCCCTCC TCCATAGTAA GTCTCATCCT
```

```
     +2 L  H  E  Y  P  V  G  S  Q  L  P  C  E  P  E  P  D  V  A  V  L  T  S  M  L  T  D
4721 CTCCACGAAT ACCCGGTAGG GTCGCAATTA CCTTGCGAGC CCGAACCGGA CGTGGCCGTG TTGACGTCCA TGCTCACTGA
     GAGGTGCTTA TGGGCCATCC CAGCGTTAAT GGAACGCTCG GGCTTGGCCT GCACCGGCAC AACTGCAGGT ACGAGTGACT
```

```
     +2   P  S  H  I  T  A  E  A  A  G  R  R  L  A  R  G  S  P  P  S  V  A  S  S  S  A
4801 TCCCTCCCAT ATAACAGCAG AGGCGGCCGG GCGAAGGTTG GCGAGGGGAT CACCCCCCTC TGTGGCCAGC TCCTCGGCTA
     AGGGAGGGTA TATTGTCGTC TCCGCCGGCC CGCTTCCAAC CGCTCCCCTA GTGGGGGGAG ACACCGGTCG AGGAGCCGAT
```

```
      +2  S   Q   L   S   A   P   S   L   K   A   T   C   T   A   N   H   D   S   P   D   A   E   L   I   E   A   N
    4881  GCCAGCTATC CGCTCCATCT CTCAAGGCAA CTTGCACCGC TAACCATGAC TCCCCTGATG CTGAGCTCAT AGAGGCCAAC
          CGGTCGATAG GCGAGGTAGA GAGTTCCGTT GAACGTGGCG ATTGGTACTG AGGGGACTAC GACTCGAGTA TCTCCGGTTG

      +2  L   L   W   R   Q   E   M   G   G   N   I   T   R   V   E   S   E   N   K   V   V   I   L   D   S   F   D
    4961  CTCCTATGGA GGCAGGAGAT GGGCGGCAAC ATCACCAGGG TTGAGTCAGA AAACAAAGTG GTGATTCTGG ACTCCTTCGA
          GAGGATACCT CCGTCCTCTA CCCGCCGTTG TAGTGGTCCC AACTCAGTCT TTTGTTTCAC CACTAAGACC TGAGGAAGCT

      +2  P   L   V   A   E   E   D   E   R   E   I   S   V   P   A   E   I   L   R   K   S   R   R   F   A   Q
    5041  TCCGCTTGTG GCGGAGGAGG ACGAGCGGGA GATCTCCGTA CCCGCAGAAA TCCTGCGGAA GTCTCGGAGA TTCGCCCAGG
          AGGCGAACAC CGCCTCCTCC TGCTCGCCCT CTAGAGGCAT GGGCGTCTTT AGGACGCCTT CAGAGCCTCT AAGCGGGTCC

      +2  A   L   P   V   W   A   R   P   D   Y   N   P   P   L   V   E   T   W   K   K   P   D   Y   E   P   P   V
    5121  CCCTGCCCGT TTGGGCGCGG CCGGACTATA ACCCCCCGCT AGTGGAGACG TGGAAAAAGC CCGACTACGA ACCACCTGTG
          GGGACGGGCA AACCCGCGCC GGCCTGATAT TGGGGGGCGA TCACCTCTGC ACCTTTTTCG GGCTGATGCT TGGTGGACAC

      +2  V   H   G   C   P   L   P   P   P   K   S   P   P   V   P   P   P   R   K   K   R   T   V   V   L   T   E
    5201  GTCCATGGCT GCCCGCTTCC ACCTCCAAAG TCCCCTCCTG TGCCTCCGCC TCGGAAGAAG CGGACGGTGG TCCTCACTGA
          CAGGTACCGA CGGGCGAAGG TGGAGGTTTC AGGGGAGGAC ACGGAGGCGG AGCCTTCTTC GCCTGCCACC AGGAGTGACT

      +2  S   T   L   S   T   A   L   A   E   L   A   T   R   S   F   G   S   S   S   T   S   G   I   T   G   D
    5281  ATCAACCCTA TCTACTGCCT TGGCCGAGCT CGCCACCAGA AGCTTTGGCA GCTCCTCAAC TTCCGGCATT ACGGGCGACA
          TAGTTGGGAT AGATGACGGA ACCGGCTCGA GCGGTGGTCT TCGAAACCGT CGAGGAGTTG AAGGCCGTAA TGCCCGCTGT

      +2  N   T   T   T   S   S   E   P   A   P   S   G   C   P   P   D   S   D   A   E   S   Y   S   S   M   P   P
    5361  ATACGACAAC ATCCTCTGAG CCCGCCCCTT CTGGCTGCCC CCCCGACTCC GACGCTGAGT CCTATTCCTC CATGCCCCCC
          TATGCTGTTG TAGGAGACTC GGGCGGGGAA GACCGACGGG GGGGCTGAGG CTGCGACTCA GGATAAGGAG GTACGGGGGG

      +2  L   E   G   E   P   G   D   P   D   L   S   D   G   S   W   S   T   V   S   S   E   A   N   A   E   D   V
                            BamHI
                            ~~~~~~~
    5441  CTGGAGGGGG AGCCTGGGGA TCCGGATCTT AGCGACGGGT CATGGTCAAC GGTCAGTAGT GAGGCCAACG CGGAGGATGT
          GACCTCCCCC TCGGACCCCT AGGCCTAGAA TCGCTGCCCA GTACCAGTTG CCAGTCATCA CTCCGGTTGC GCCTCCTACA

      +2  V   C   C   S   M   S   Y   S   W   T   G   A   L   V   T   P   C   A   A   E   E   Q   K   L   P   I
    5521  CGTGTGCTGC TCAATGTCTT ACTCTTGGAC AGGCGCACTC GTCACCCCGT GCGCCGCGGA AGAACAGAAA CTGCCCATCA
          GCACACGACG AGTTACAGAA TGAGAACCTG TCCGCGTGAG CAGTGGGGCA CGCGGCGCCT TCTTGTCTTT GACGGGTAGT

      +2  N   A   L   S   N   S   L   L   R   H   H   N   L   V   Y   S   T   T   S   R   S   A   C   Q   R   Q   K
    5601  ATGCACTAAG CAACTCGTTG CTACGTCACC ACAATTTGGT GTATTCCACC ACCTCACGCA GTGCTTGCCA AAGGCAGAAG
          TACGTGATTC GTTGAGCAAC GATGCAGTGG TGTTAAACCA CATAAGGTGG TGGAGTGCGT CACGAACGGT TTCCGTCTTC

      +2  K   V   T   F   D   R   L   Q   V   L   D   S   H   Y   Q   D   V   L   K   E   V   K   A   A   S   K
    5681  AAAGTCACAT TTGACAGACT GCAAGTTCTG GACAGCCATT ACCAGGACGT ACTCAAGGAG GTTAAAGCAG CGGCGTCAAA
          TTTCAGTGTA AACTGTCTGA CGTTCAAGAC CTGTCGGTAA TGGTCCTGCA TGAGTTCCTC CAATTTCGTC GCCGCAGTTT

      +2  V   K   A   N   L   L   S   V   E   E   A   C   S   L   T   P   P   H   S   A   K   S   K   F   G   Y
    5761  AGTGAAGGCT AACTTGCTAT CCGTAGAGGA AGCTTGCAGC CTGACGCCCC CACACTCAGC CAAATCCAAG TTTGGTTATG
          TCACTTCCGA TTGAACGATA GGCATCTCCT TCGAACGTCG GACTGCGGGG GTGTGAGTCG GTTTAGGTTC AAACCAATAC

      +2  G   A   K   D   V   R   C   H   A   R   K   A   V   T   H   I   N   S   V   W   K   D   L   L   E   D   N
    5841  GGGCAAAAGA CGTCCGTTGC CATGCCAGAA AGGCCGTAAC CCACATCAAC TCCGTGTGGA AAGACCTTCT GGAAGACAAT
          CCCGTTTTCT GCAGGCAACG GTACGGTCTT TCCGGCATTG GGTGTAGTTG AGGCACACCT TTCTGGAAGA CCTTCTGTTA

      +2  V   T   P   I   D   T   T   I   M   A   K   N   E   V   F   C   V   Q   P   E   K   G   G   R   K   P   A
    5921  GTAACACCAA TAGACACTAC CATCATGGCT AAGAACGAGG TTTTCTGCGT TCAGCCTGAG AAGGGGGGTC GTAAGCCAGC
          CATTGTGGTT ATCTGTGATG GTAGTACCGA TTCTTGCTCC AAAAGACGCA AGTCGGACTC TTCCCCCCAG CATTCGGTCG
```

```
     +2    R   L   I   V   F   P   D   L   G   V   R   V   C   E   K   M   A   L   Y   D   V   V   T   K   L   P
6001  TCGTCTCATC GTGTTCCCCG ATCTGGGCGT GCGCGTGTGC GAAAAGATGG CTTTGTACGA CGTGGTTACA AAGCTCCCCT
      AGCAGAGTAG CACAAGGGGC TAGACCCGCA CGCGCACACG CTTTTCTACC GAAACATGCT GCACCAATGT TTCGAGGGGA
```

```
     +2    L   A   V   M   G   S   S   Y   G   F   Q   Y   S   P   G   Q   R   V   E   F   L   V   Q   A   W   K   S
                                                                              EcoRI
                                                                              ~~~~~~
6081  TGGCCGTGAT GGGAAGCTCC TACGGATTCC AATACTCACC AGGACAGCGG GTTGAATTCC TCGTGCAAGC GTGGAAGTCC
      ACCGGCACTA CCCTTCGAGG ATGCCTAAGG TTATGAGTGG TCCTGTCGCC CAACTTAAGG AGCACGTTCG CACCTTCAGG
```

```
     +2    K   K   T   P   M   G   F   S   Y   D   T   R   C   F   D   S   T   V   T   E   S   D   I   R   T   E   E
6161  AAGAAAACCC CAATGGGGTT CTCGTATGAT ACCCGCTGCT TTGACTCCAC AGTCACTGAG AGCGACATCC GTACGGAGGA
      TTCTTTTGGG GTTACCCCAA GAGCATACTA TGGGCGACGA AACTGAGGTG TCAGTGACTC TCGCTGTAGG CATGCCTCCT
```

```
     +2    A   I   Y   Q   C   C   D   L   D   P   Q   A   R   V   A   I   K   S   L   T   E   R   L   Y   V   G
6241  GGCAATCTAC CAATGTTGTG ACCTCGACCC CCAAGCCCGC GTGGCCATCA AGTCCCTCAC CGAGAGGCTT TATGTTGGGG
      CCGTTAGATG GTTACAACAC TGGAGCTGGG GGTTCGGGCG CACCGGTAGT TCAGGGAGTG GCTCTCCGAA ATACAACCCC
```

```
     +2    G   P   L   T   N   S   R   G   E   N   C   G   Y   R   R   C   R   A   S   G   V   L   T   T   S   C   G
6321  GCCCTCTTAC CAATTCAAGG GGGGAGAACT GCGGCTATCG CAGGTGCCGC GCGAGCGGCG TACTGACAAC TAGCTGTGGT
      CGGGAGAATG GTTAAGTTCC CCCCTCTTGA CGCCGATAGC GTCCACGGCG CGCTCGCCGC ATGACTGTTG ATCGACACCA
```

```
     +2    N   T   L   T   C   Y   I   K   A   R   A   A   C   R   A   A   G   L   Q   D   C   T   M   L   V   C   G
6401  AACACCCTCA CTTGCTACAT CAAGGCCCGG GCAGCCTGTC GAGCCGCAGG GCTCCAGGAC TGCACCATGC TCGTGTGTGG
      TTGTGGGAGT GAACGATGTA GTTCCGGGCC CGTCGGACAG CTCGGCGTCC CGAGGTCCTG ACGTGGTACG AGCACACACC
```

```
     +2    D   D   L   V   V   I   C   E   S   A   G   V   Q   E   D   A   A   S   L   R   A   F   T   E   A   M
6481  CGACGACTTA GTCGTTATCT GTGAAAGCGC GGGGGTCCAG GAGGACGCGG CGAGCCTGAG AGCCTTCACG GAGGCTATGA
      GCTGCTGAAT CAGCAATAGA CACTTTCGCG CCCCCAGGTC CTCCTGCGCC GCTCGGACTC TCGGAAGTGC CTCCGATACT
```

```
     +2    T   R   Y   S   A   P   P   G   D   P   P   Q   P   E   Y   D   L   E   L   I   T   S   C   S   S   N   V
6561  CCAGGTACTC CGCCCCCCCT GGGGACCCCC CACAACCAGA ATACGACTTG GAGCTCATAA CATCATGCTC CTCCAACGTG
      GGTCCATGAG GCGGGGGGGA CCCCTGGGGG GTGTTGGTCT TATGCTGAAC CTCGAGTATT GTAGTACGAG GAGGTTGCAC
```

```
     +2    S   V   A   H   D   G   A   G   K   R   V   Y   Y   L   T   R   D   P   T   T   P   L   A   R   A   A   W
6641  TCAGTCGCCC ACGACGGCGC TGGAAAGAGG GTCTACTACC TCACCCGTGA CCCTACAACC CCCCTCGCGA GAGCTGCGTG
      AGTCAGCGGG TGCTGCCGCG ACCTTTCTCC CAGATGATGG AGTGGGCACT GGGATGTTGG GGGGAGCGCT CTCGACGCAC
```

```
     +2    E   T   A   R   H   T   P   V   N   S   W   L   G   N   I   I   M   F   A   P   T   L   W   A   R   M
6721  GGAGACAGCA AGACACACTC CAGTCAATTC CTGGCTAGGC AACATAATCA TGTTTGCCCC CACACTGTGG GCGAGGATGA
      CCTCTGTCGT TCTGTGTGAG GTCAGTTAAG GACCGATCCG TTGTATTAGT ACAAACGGGG GTGTGACACC CGCTCCTACT
```

```
     +2    I   L   M   T   H   F   F   S   V   L   I   A   R   D   Q   L   E   Q   A   L   D   C   E   I   Y   G   A
6801  TACTGATGAC CCATTTCTTT AGCGTCCTTA TAGCCAGGGA CCAGCTTGAA CAGGCCCTCG ATTGCGAGAT CTACGGGGCC
      ATGACTACTG GGTAAAGAAA TCGCAGGAAT ATCGGTCCCT GGTCGAACTT GTCCGGGAGC TAACGCTCTA GATGCCCCGG
```

```
     +2    C   Y   S   I   E   P   L   D   L   P   P   I   I   Q   R   L   H   G   L   S   A   F   S   L   H   S   Y
6881  TGCTACTCCA TAGAACCACT GGATCTACCT CCAATCATTC AAAGACTCCA TGGCCTCAGC GCATTTTCAC TCCACAGTTA
      ACGATGAGGT ATCTTGGTGA CCTAGATGGA GGTTAGTAAG TTTCTGAGGT ACCGGAGTCG CGTAAAAGTG AGGTGTCAAT
```

```
     +2    S   P   G   E   I   N   R   V   A   A   C   L   R   K   L   G   V   P   P   L   R   A   W   R   H   R
6961  CTCTCCAGGT GAAATCAATA GGGTGGCCGC ATGCCTCAGA AAACTTGGGG TACCGCCCTT GCGAGCTTGG AGACACCGGG
      GAGAGGTCCA CTTTAGTTAT CCCACCGGCG TACGGAGTCT TTTGAACCCC ATGGCGGGAA CGCTCGAACC TCTGTGGCCC
```

```
     +2    A   R   S   V   R   A   R   L   L   A   R   G   G   R   A   A   I   C   G   K   Y   L   F   N   W   A   V
7041  CCCGGAGCGT CCGCGCTAGG CTTCTGGCCA GAGGAGGCAG GGCTGCCATA TGTGGCAAGT ACCTCTTCAA CTGGGCAGTA
      GGGCCTCGCA GGCGCGATCC GAAGACCGGT CTCCTCCGTC CCGACGGTAT ACACCGTTCA TGGAGAAGTT GACCCGTCAT
```

```
     +2  R  T  K  L    K  L  T    P  I  A    A  A  G  Q    L  D  L    S  G  W    F  T  A  G    Y  S  G
7121 AGAACAAAGC TCAAACTCAC TCCAATAGCG GCCGCTGGCC AGCTGGACTT GTCCGGCTGG TTCACGGCTG GCTACAGCGG
     TCTTGTTTCG AGTTTGAGTG AGGTTATCGC CGGCGACCGG TCGACCTGAA CAGGCCGACC AAGTGCCGAC CGATGTCGCC


     +2  G  D  I  Y  H  S  V    S  H  A    R  P  R    W  I  W  F    C  L  L    L  L  A    A  G  V
7201 GGGAGACATT TATCACAGCG TGTCTCATGC CCGGCCCCGC TGGATCTGGT TTTGCCTACT CCTGCTTGCT GCAGGGGTAG
     CCCTCTGTAA ATAGTGTCGC ACAGAGTACG GGCCGGGGCG ACCTAGACCA AAACGGATGA GGACGAACGA CGTCCCCATC


     +2 G  I  Y  L    L  P  N    R
7281 GCATCTACCT CCTCCCCAAC CGATGAAGGT TGGGGTAAAC ACTCCGGCCT AAAAAAAAAA AAAAATCTAG AAAGGCGCGC
     CGTAGATGGA GGAGGGGTTG GCTACTTCCA ACCCCATTTG TGAGGCCGGA TTTTTTTTTT TTTTTAGATC TTTCCGCGCG


                     BamHI      MluI
                     ~~~~~~     ~~~~~~
7361 CAAGATATCA AGGATCCACT ACGCGTTAGA GCTCGCTGAT CAGCCTCGAC TGTGCCTTCT AGTTGCCAGC CATCTGTTGT
     GTTCTATAGT TCCTAGGTGA TGCGCAATCT CGAGCGACTA GTCGGAGCTG ACACGGAAGA TCAACGGTCG GTAGACAACA


7441 TTGCCCCTCC CCCGTGCCTT CCTTGACCCT GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAAATGAG GAAATTGCAT
     AACGGGGAGG GGGCACGGAA GGAACTGGGA CCTTCCACGG TGAGGGTGAC AGGAAAGGAT TATTTTACTC CTTTAACGTA


7521 CGCATTGTCT GAGTAGGTGT CATTCTATTC TGGGGGGTGG GGTGGGGCAG GACAGCAAGG GGGAGGATTG GGAAGACAAT
     GCGTAACAGA CTCATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC CTGTCGTTCC CCCTCCTAAC CCTTCTGTTA


7601 AGCAGGCATG CTGGGGAGCT CTTCCGCTTC CTCGCTCACT GACTCGCTGC GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT
     TCGTCCGTAC GACCCCTCGA GAAGGCGAAG GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCCGACGCC GCTCGCCATA


7681 CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG
     GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCTATTG CGTCCTTTCT TGTACACTCG TTTTCCGGTC


7761 CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA
     GTTTTCCGGT CCTTGGCATT TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCGGGG GGACTGCTCG TAGTGTTTTT


7841 TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC
     AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCCTGAT ATTTCTATGG TCCGCAAAGG GGGACCTTCG AGGGAGCACG


7921 GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCAATGC
     CGAGAGGACA AGGCTGGGAC GGCGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCGA AAGAGTTACG


8001 TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA
     AGTGCGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GGTTCGACCC GACACACGTG CTTGGGGGGC AAGTCGGGCT


8081 CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA GCAGCCACTG
     GGCGACGCGG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CGGTGACCGT CGTCGGTGAC


8161 GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACTACGG CTACACTAGA
     CATTGTCCTA ATCGTCTCGC TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT


8241 AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAACA
     TCCTGTCATA AACCATAGAC GCGAGACGAC TTCGGTCAAT GGAAGCCTTT TTCTCAACCA TCGAGAACTA GGCCGTTTGT


8321 AACCACCGCT GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT
     TTGGTGGCGA CCATCGCCAC CAAAAAAACA AACGTTCGTC GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTTCTAGGAA


8401 TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG
     ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT CCCTAAAACC AGTACTCTAA TAGTTTTTCC
```

```
8481   ATCTTCACCT AGATCCTTTT AAAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT GGTCTGACAG
       TAGAAGTGGA TCTAGGAAAA TTTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA CCAGACTGTC

8561   TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC TGTCTATTTC GTTCATCCAT AGTTGCCTGA CTCCCCGTCG
       AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGGGCAGC

8641   TGTAGATAAC TACGATACGG GAGGGCTTAC CATCTGGCCC CAGTGCTGCA ATGATACCGC GAGACCCACG CTCACCGGCT
       ACATCTATTG ATGCTATGCC CTCCCGAATG GTAGACCGGG GTCACGACGT TACTATGGCG CTCTGGGTGC GAGTGGCCGA

8721   CCAGATTTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCCG AGCGCAGAAG TGGTCCTGCA ACTTTATCCG CCTCCATCCA
       GGTCTAAATA GTCGTTATTT GGTCGGTCGG CCTTCCCGGC TCGCGTCTTC ACCAGGACGT TGAAATAGGC GGAGGTAGGT

8801   GTCTATTAAT TGTTGCCGGG AAGCTAGAGT AAGTAGTTCG CCAGTTAATA GTTTGCGCAA CGTTGTTGCC ATTGCTACAG
       CAGATAATTA ACAACGGCCC TTCGATCTCA TTCATCAAGC GGTCAATTAT CAAACGCGTT GCAACAACGG TAACGATGTC

8881   GCATCGTGGT GTCACGCTCG TCGTTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGCGAGT TACATGATCC
       CGTAGCACCA CAGTGCGAGC AGCAAACCAT ACCGAAGTAA GTCGAGGCCA AGGGTTGCTA GTTCCGCTCA ATGTACTAGG

8961   CCCATGTTGT GCAAAAAAGC GGTTAGCTCC TTCGGTCCTC CGATCGTTGT CAGAAGTAAG TTGGCCGCAG TGTTATCACT
       GGGTACAACA CGTTTTTTCG CCAATCGAGG AAGCCAGGAG GCTAGCAACA GTCTTCATTC AACCGGCGTC ACAATAGTGA

9041   CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCATG CCATCCGTAA GATGCTTTTC TGTGACTGGT GAGTACTCAA
       GTACCAATAC CGTCGTGACG TATTAAGAGA ATGACAGTAC GGTAGGCATT CTACGAAAAG ACACTGACCA CTCATGAGTT

9121   CCAAGTCATT CTGAGAATAG TGTATGCGGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC GGGATAATAC CGCGCCACAT
       GGTTCAGTAA GACTCTTATC ACATACGCCG CTGGCTCAAC GAGAACGGGC CGCAGTTATG CCCTATTATG GCGCGGTGTA

9201   AGCAGAACTT TAAAAGTGCT CATCATTGGA AAACGTTCTT CGGGGCGAAA ACTCTCAAGG ATCTTACCGC TGTTGAGATC
       TCGTCTTGAA ATTTTCACGA GTAGTAACCT TTTGCAAGAA GCCCCGCTTT TGAGAGTTCC TAGAATGGCG ACAACTCTAG

9281   CAGTTCGATG TAACCCACTC GTGCACCCAA CTGATCTTCA GCATCTTTTA CTTTCACCAG CGTTTCTGGG TGAGCAAAAA
       GTCAAGCTAC ATTGGGTGAG CACGTGGGTT GACTAGAAGT CGTAGAAAAT GAAAGTGGTC GCAAAGACCC ACTCGTTTTT

9361   CAGGAAGGCA AAATGCCGCA AAAAAGGGAA TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTTCAATAT
       GTCCTTCCGT TTTACGGCGT TTTTTCCCTT ATTCCCGCTG TGCCTTTACA ACTTATGAGT ATGAGAAGGA AAAAGTTATA

9441   TATTGAAGCA TTTATCAGGG TTATTGTCTC ATGAGCGGAT ACATATTTGA ATGTATTTAG AAAAATAAAC AAATAGGGGT
       ATAACTTCGT AAATAGTCCC AATAACAGAG TACTCGCCTA TGTATAAACT TACATAAATC TTTTTATTTG TTTATCCCCA

9521   TCCGCGCACA TTTCCCCGAA AAGTGCCACC TGACGTCTAA GAAACCATTA TTATCATGAC ATTAACCTAT AAAAATAGGC
       AGGCGCGTGT AAAGGGGCTT TTCACGGTGG ACTGCAGATT CTTTGGTAAT AATAGTACTG TAATTGGATA TTTTTATCCG

9601   GTATCACGAG GCCCTTTCGT C
       CATAGTGCTC CGGGAAAGCA G
```
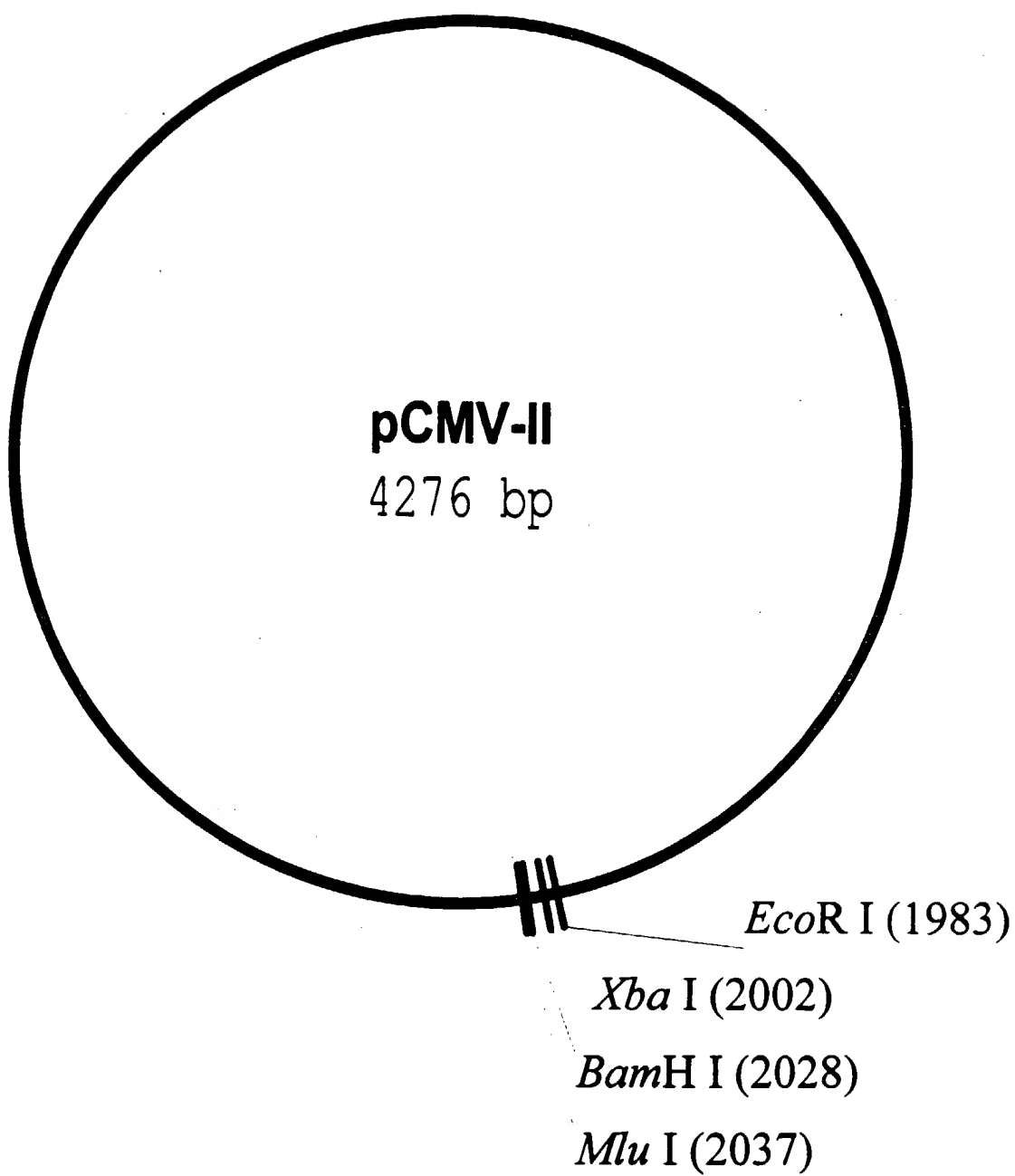
FIGURE 6

pCMV-II
4276 bp

EcoR I (1983)

Xba I (2002)

BamH I (2028)

Mlu I (2037)

```
   1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT
      AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTCGCCTA

  81  GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA
      CGGCCCTCGT CTGTTCGGGC AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC GCCGTAGTCT

 161  GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG
      CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC

 241  AATAGCTCAG AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGGCGGA
      TTATCGAGTC TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT ACCCCGCCTC TTACCCGCCT

 321  ACTGGGCGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
      TGACCCGCCC CTCCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

 401  CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT
      GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA

 481  AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCCATT
      TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA

 561  GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT
      CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

 641  AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG TCAATGACGG TAAATGGCCC
      TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC AGTTACTGCC ATTTACGGG

 721  GCCTGGCATT ATGCCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
      CGGACCGTAA TACGGGTCAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

 801  CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA
      GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT

 881  TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC CCCGTTGACG
      AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

 961  CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG
      GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1041  CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG GGAACGGTGC ATTGGAACGC
      GGTAGGTGCG ACAAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCCGGC CCTTGCCACG TAACCTTGCG

1121  GGATTCCCCG TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
      CCTAAGGGGC ACGGTTCTCA CTGCATTCAT GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201  CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA
      GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

1281  TTGACCATTA TTGACCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACTAT
      AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

1361  CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT
      GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA

1441  TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT
      ATAAATGTTT AAGTGTATAT GTTGTTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT ATCGCACCCT AGAGGCTGTA
```

```
1521  CTCGGGTACG TGTTCCGGAC ATGGGCTCTT CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
      GAGCCCATGC ACAAGGCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601  GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC
      CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTTACGG GTGGTGGTGG

1681  AGTGTGCCGC ACAAGGCCGT GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT
      TCACACGGCG TGTTCCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA CCTGCGTCTA

1761  GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT
      CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841  TGCGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT
      ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

                                                                  EcoRI
                                                                  ~~~~~~
1921  GACAGACTAA CAGACTGTTC CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCAGA CTCGAGCAAG
      CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTCT GAGCTCGTTC

      XbaI                        BamHI      MluI
      ~~~~~~                      ~~~~~~~    ~~~~~~~
2001  TCTAGAAAGG CGCGCCAAGA TATCAAGGAT CCACTACGCG TTAGAGCTCG CTGATCAGCC TCGACTGTGC CTTCTAGTTG
      AGATCTTTCC GCGCGGTTCT ATAGTTCCTA GGTGATGCGC AATCTCGAGC GACTAGTCGG AGCTGACACG GAAGATCAAC

2081  CCAGCCATCT GTTGTTTGCC CCTCCCCCGT GCCTTCCTTG ACCCTGGAAG GTGCCACTCC CACTGTCCTT TCCTAATAAA
      GGTCGGTAGA CAACAAACGG GGAGGGGGCA CGGAAGGAAC TGGGACCTTC CACGGTGAGG GTGACAGGAA AGGATTATTT

2161  ATGAGGAAAT TGCATCGCAT TGTCTGAGTA GGTGTCATTC TATTCTGGGG GGTGGGGTGG GGCAGGACAG CAAGGGGGAG
      TACTCCTTTA ACGTAGCGTA ACAGACTCAT CCACAGTAAG ATAAGACCCC CCACCCCACC CCGTCCTGTC GTTCCCCCTC

2241  GATTGGGAAG ACAATAGCAG GCATGCTGGG GAGCTCTTCC GCTTCCTCGC TCACTGACTC GCTGCGCTCG GTCGTTCGGC
      CTAACCCTTC TGTTATCGTC CGTACGACCC CTCGAGAAGG CGAAGGAGCG AGTGACTGAG CGACGCGAGC CAGCAAGCCG

2321  TGCGGCGAGC GGTATCAGCT CACTCAAAGG CGGTAATACG GTTATCCACA GAATCAGGGG ATAACGCAGG AAAGAACATG
      ACGCCGCTCG CCATAGTCGA GTGAGTTTCC GCCATTATGC CAATAGGTGT CTTAGTCCCC TATTGCGTCC TTTCTTGTAC

2401  TGAGCAAAAG GCCAGCAAAA GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT GGCGTTTTTC CATAGGCTCC GCCCCCCTGA
      ACTCGTTTTC CGGTCGTTTT CCGGTCCTTG GCATTTTTCC GGCGCAACGA CCGCAAAAAG GTATCCGAGG CGGGGGGACT

2481  CGAGCATCAC AAAAATCGAC GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAAG ATACCAGGCG TTTCCCCCTG
      GCTCGTAGTG TTTTTAGCTG CGAGTTCAGT CTCCACCGCT TTGGGCTGTC CTGATATTTC TATGGTCCGC AAAGGGGGAC

2561  GAAGCTCCCT CGTGCGCTCT CCTGTTCCGA CCCTGCCGCT TACCGGATAC CTGTCCGCCT TTCTCCCTTC GGGAAGCGTG
      CTTCGAGGGA GCACGCGAGA GGACAAGGCT GGGACGGCGA ATGGCCTATG GACAGGCGGA AAGAGGGAAG CCCTTCGCAC

2641  GCGCTTTCTC AATGCTCACG CTGTAGGTAT CTCAGTTCGG TGTAGGTCGT TCGCTCCAAG CTGGGCTGTG TGCACGAACC
      CGCGAAAGAG TTACGAGTGC GACATCCATA GAGTCAAGCC ACATCCAGCA AGCGAGGTTC GACCCGACAC ACGTGCTTGG

2721  CCCCGTTCAG CCCGACCGCT GCGCCTTATC CGGTAACTAT CGTCTTGAGT CCAACCCGGT AAGACACGAC TTATCGCCAC
      GGGGCAAGTC GGGCTGGCGA CGCGGAATAG GCCATTGATA GCAGAACTCA GGTTGGGCCA TTCTGTGCTG AATAGCGGTG

2801  TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTAGGCGGT GCTACAGAGT TCTTGAAGTG GTGGCCTAAC
      ACCGTCGTCG GTGACCATTG TCCTAATCGT CTCGCTCCAT ACATCCGCCA CGATGTCTCA AGAACTTCAC CACCGGATTG

2881  TACGGCTACA CTAGAAGGAC AGTATTTGGT ATCTGCGCTC TGCTGAAGCC AGTTACCTTC GGAAAAAGAG TTGGTAGCTC
      ATGCCGATGT GATCTTCCTG TCATAAACCA TAGACGCGAG ACGACTTCGG TCAATGGAAG CCTTTTTCTC AACCATCGAG
```

```
2961  TTGATCCGGC AAACAAACCA CCGCTGGTAG CGGTGGTTTT TTTGTTTGCA AGCAGCAGAT TACGCGCAGA AAAAAAGGAT
      AACTAGGCCG TTTGTTTGGT GGCGACCATC GCCACCAAAA AAACAAACGT TCGTCGTCTA ATGCGCGTCT TTTTTTCCTA

3041  CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACTCAC GTTAAGGGAT TTTGGTCATG
      GAGTTCTTCT AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTTGAGTG CAATTCCCTA AAACCAGTAC

3121  AGATTATCAA AAAGGATCTT CACCTAGATC CTTTTAAATT AAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA
      TCTAATAGTT TTTCCTAGAA GTGGATCTAG GAAAATTTAA TTTTTACTTC AAAATTTAGT TAGATTTCAT ATATACTCAT

3201  AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTCAG CGATCTGTCT ATTTCGTTCA TCCATAGTTG
      TTGAACCAGA CTGTCAATGG TTACGAATTA GTCACTCCGT GGATAGAGTC GCTAGACAGA TAAAGCAAGT AGGTATCAAC

3281  CCTGACTCCC CGTCGTGTAG ATAACTACGA TACGGGAGGG CTTACCATCT GGCCCCAGTG CTGCAATGAT ACCGCGAGAC
      GGACTGAGGG GCAGCACATC TATTGATGCT ATGCCCTCCC GAATGGTAGA CCGGGGTCAC GACGTTACTA TGGCGCTCTG

3361  CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC CAGCCGGAAG GGCCGAGCGC AGAAGTGGTC CTGCAACTTT
      GGTGCGAGTG GCCGAGGTCT AAATAGTCGT TATTTGGTCG GTCGGCCTTC CCGGCTCGCG TCTTCACCAG GACGTTGAAA

3441  ATCCGCCTCC ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT TAATAGTTTG CGCAACGTTG
      TAGGCGGAGG TAGGTCAGAT AATTAACAAC GGCCCTTCGA TCTCATTCAT CAAGCGGTCA ATTATCAAAC GCGTTGCAAC

3521  TTGCCATTGC TACAGGCATC GTGGTGTCAC GCTCGTCGTT TGGTATGGCT TCATTCAGCT CCGGTTCCCA ACGATCAAGG
      AACGGTAACG ATGTCCGTAG CACCACAGTG CGAGCAGCAA ACCATACCGA AGTAAGTCGA GGCCAAGGGT TGCTAGTTCC

3601  CGAGTTACAT GATCCCCCAT GTTGTGCAAA AAAGCGGTTA GCTCCTTCGG TCCTCCGATC GTTGTCAGAA GTAAGTTGGC
      GCTCAATGTA CTAGGGGGTA CAACACGTTT TTTCGCCAAT CGAGGAAGCC AGGAGGCTAG CAACAGTCTT CATTCAACCG

3681  CGCAGTGTTA TCACTCATGG TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC TTTTCTGTGA
      GCGTCACAAT AGTGAGTACC AATACCGTCG TGACGTATTA AGAGAATGAC AGTACGGTAG GCATTCTACG AAAAGACACT

3761  CTGGTGAGTA CTCAACCAAG TCATTCTGAG AATAGTGTAT GCGGCGACCG AGTTGCTCTT GCCCGGCGTC AATACGGGAT
      GACCACTCAT GAGTTGGTTC AGTAAGACTC TTATCACATA CGCCGCTGGC TCAACGAGAA CGGGCCGCAG TTATGCCCTA

3841  AATACCGCGC CACATAGCAG AACTTTAAAA GTGCTCATCA TTGGAAAACG TTCTTCGGGG CGAAAACTCT CAAGGATCTT
      TTATGGCGCG GTGTATCGTC TTGAAATTTT CACGAGTAGT AACCTTTTGC AAGAAGCCCC GCTTTTGAGA GTTCCTAGAA

3921  ACCGCTGTTG AGATCCAGTT CGATGTAACC CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTTT
      TGGCGACAAC TCTAGGTCAA GCTACATTGG GTGAGCACGT GGGTTGACTA GAAGTCGTAG AAAATGAAAG TGGTCGCAAA

4001  CTGGGTGAGC AAAAACAGGA AGGCAAAATG CCGCAAAAAA GGGAATAAGG GCGACACGGA AATGTTGAAT ACTCATACTC
      GACCCACTCG TTTTTGTCCT TCCGTTTTAC GGCGTTTTTT CCCTTATTCC CGCTGTGCCT TTACAACTTA TGAGTATGAG

4081  TTCCTTTTTC AATATTATTG AAGCATTTAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA TTTAGAAAAA
      AAGGAAAAAG TTATAATAAC TTCGTAAATA GTCCCAATAA CAGAGTACTC GCCTATGTAT AAACTTACAT AAATCTTTTT

4161  TAAACAAATA GGGGTTCCGC GCACATTTCC CCGAAAAGTG CCACCTGACG TCTAAGAAAC CATTATTATC ATGACATTAA
      ATTTGTTTAT CCCCAAGGCG CGTGTAAAGG GGCTTTTCAC GGTGGACTGC AGATTCTTTG GTAATAATAG TACTGTAATT

4241  CCTATAAAAA TAGGCGTATC ACGAGGCCCT TTCGTC
      GGATATTTTT ATCCGCATAG TGCTCCGGGA AAGCAG
```

FIGURE 8

Stu I (211)

pCMV-NS34A
6300 bp

*Eco*R I (1983)

Mlu I (4061)

*Bam*H I (4052)

Stu I (3607)

NS34A ORF
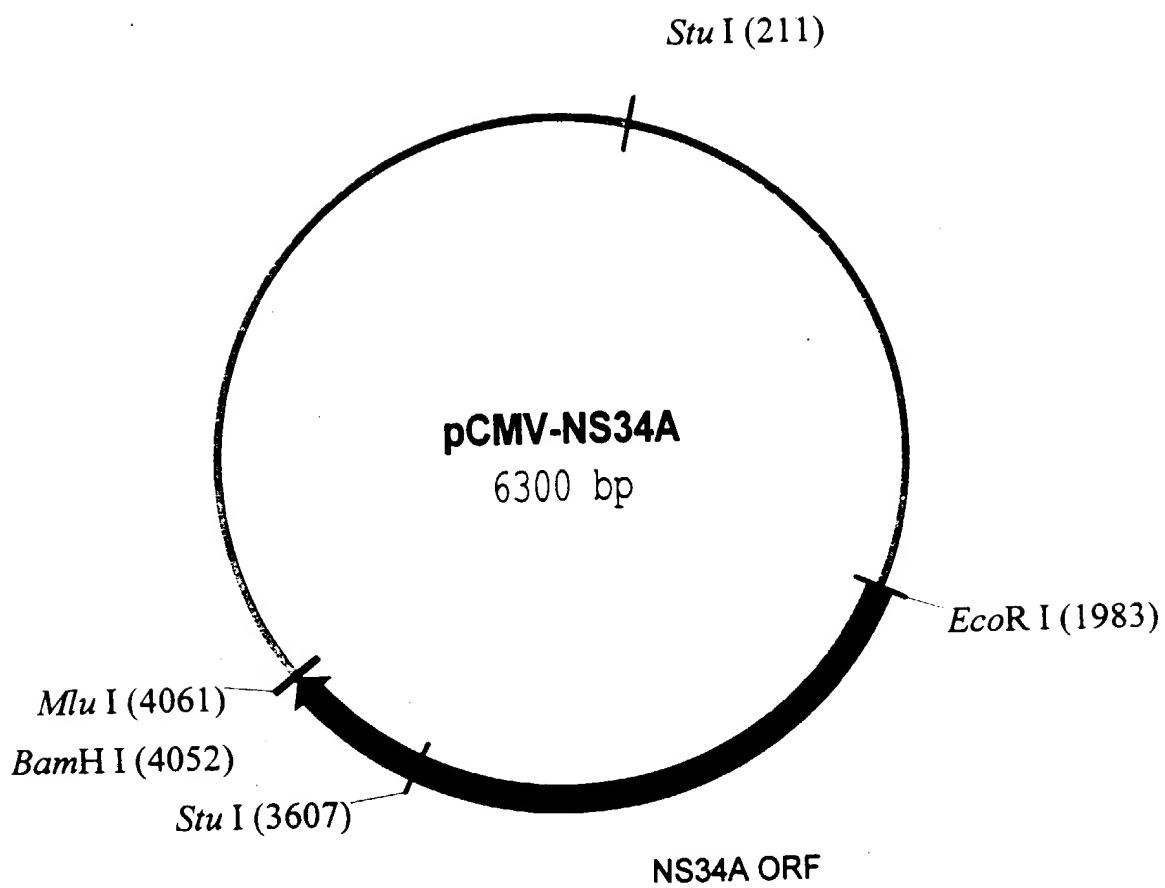
FIGURE 9 - Page 1

```
  1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
     AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

 51  GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
     CTCTGCCAGT GTCGAACAGA CATTCGCCTA CGGCCCTCGT CTGTTCGGGC

101  TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
     AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

151  CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA
     GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT

               StuI
               ~~~~~~~
201  AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG
     TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

251  AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA
     TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT

301  TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG
     ACCCCGCCTC TTACCCGCCT TGACCCGCCC CTCCCTTAAT AACCGATAAC

351  GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
     CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401  CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA
     GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

451  TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCCG
     ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

501  CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
     GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

551  CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA
     GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

601  GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA
     CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

651  CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG
     GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

701  TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA
     AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT

751  CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
     GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

801  CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG
     GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC

851  ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG
     TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC
```

```
 901   TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC
       AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG

 951   CCCGTTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA
       GGGCAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT

1001   GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC
       CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC GGTAGGTGCG

1051   TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG
       ACAAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCCGGC

1101   GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT GACGTAAGTA
       CCTTGCCACG TAACCTTGCG CCTAAGGGGC ACGGTTCTCA CTGCATTCAT

1151   CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
       GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201   CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA
       GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT

1251   TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA TTGACCACTC
       ACCATATCGA ATCGGATATC CACACCCAAT AACTGGTAAT AACTGGTGAG

1301   CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG
       GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC

1351   CCACAACTAT CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT
       GGTGTTGATA GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA

1401   GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT TATTTACAAA
       CTGTGCCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA ATAAATGTTT

1451   TTCACATATA CAACAACGCC GTCCCCGTG CCCGCAGTTT TTATTAAACA
       AAGTGTATAT GTTGTTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT

1501   TAGCGTGGGA TCTCCGACAT CTCGGGTACG TGTTCCGGAC ATGGGCTCTT
       ATCGCACCCT AGAGGCTGTA GAGCCCATGC ACAAGGCCTG TACCCGAGAA

1551   CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
       GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601   GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA
       CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT

1651   CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCGC ACAAGGCCGT
       GAATCCGTGT CGTGTTACGG GTGGTGGTGG TCACACGGCG TGTTCCGGCA

1701   GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT
       CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA

1751   GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT
       CCTGCGTCTA CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA

1801   GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TGCGGTGCTG
       CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC
```

```
1851  TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG
      AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC
```

```
1901  CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG
      GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC
```

```
      +2                                             M  A  P
                                        EcoRI
                                        ~~~~~~
1951  GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCACC ATGGCGCCCA
      CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTGG TACCGCGGGT
```

```
      +2 I  T  A  Y  A  Q  Q   T  R  G  L   L  G  C   I  I  T
2001  TCACGGCGTA CGCCCAGCAG ACAAGGGGCC TCCTAGGGTG CATAATCACC
      AGTGCCGCAT GCGGGTCGTC TGTTCCCCGG AGGATCCCAC GTATTAGTGG
```

```
      +2 S  L  T  G  R  D  K  N  Q  V   E  G  E  V   Q  I  V
2051  AGCCTAACTG GCCGGGACAA AAACCAAGTG GAGGGTGAGG TCCAGATTGT
      TCGGATTGAC CGGCCCTGTT TTTGGTTCAC CTCCCACTCC AGGTCTAACA
```

```
      +2  S  T  A  A  Q  T  F  L  A  T   C  I  N  G  V  C
2101  GTCAACTGCT GCCCAAACCT TCCTGGCAAC GTGCATCAAT GGGGTGTGCT
      CAGTTGACGA CGGGTTTGGA AGGACCGTTG CACGTAGTTA CCCCACACGA
```

```
      +2 W  T  V  Y  H  G  A   G  T  R  T   I  A  S   P  K  G
2151  GGACTGTCTA CCACGGGGCC GGAACGAGGA CCATCGCGTC ACCCAAGGGT
      CCTGACAGAT GGTGCCCCGG CCTTGCTCCT GGTAGCGCAG TGGGTTCCCA
```

```
      +2  P  V  I  Q  M  Y  T  N  V  D   Q  D  L  V   G  W  P
2201  CCTGTCATCC AGATGTATAC CAATGTAGAC CAAGACCTTG TGGGCTGGCC
      GGACAGTAGG TCTACATATG GTTACATCTG GTTCTGGAAC ACCCGACCGG
```

```
      +2  A  S  Q  G  T  R  S  L  T  P   C  T  C   G  S  S
2251  CGCTTCGCAA GGTACCCGCT CATTGACACC CTGCACTTGC GGCTCCTCGG
      GCGAAGCGTT CCATGGGCGA GTAACTGTGG GACGTGAACG CCGAGGAGCC
```

```
      +2 D  L  Y  L   V  T  R  H  A  D  V   I  P  V   R  R  R
2301  ACCTTTACCT GGTCACGAGG CACGCCGATG TCATTCCCGT GCGCCGGCGG
      TGGAAATGGA CCAGTGCTCC GTGCGGCTAC AGTAAGGGCA CGCGGCCGCC
```

```
      +2 G  D  S  R  G  S  L   L  S  P  R  P  I  S   Y  L  K
2351  GGTGATAGCA GGGGCAGCCT GCTGTCGCCC CGGCCCATTT CCTACTTGAA
      CCACTATCGT CCCCGTCGGA CGACAGCGGG GCCGGGTAAA GGATGAACTT
```

```
      +2  G  S  S  G  G  P  L  L  C  P  A  G  H   A  V  G
2401  AGGCTCCTCG GGGGGTCCGC TGTTGTGCCC CGCGGGGCAC GCCGTGGGCA
      TCCGAGGAGC CCCCCAGGCG ACAACACGGG GCGCCCCGTG CGGCACCCGT
```

```
      +2 I  F  R  A  A  V  C   T  R  G  V   A  K  A   V  D  F
2451  TATTTAGGGC CGCGGTGTGC ACCCGTGGAG TGGCTAAGGC GGTGGACTTT
      ATAAATCCCG GCGCCACACG TGGGCACCTC ACCGATTCCG CCACCTGAAA
```

```
      +2 I  P  V  E  N  L  E   T  T  M  R  S  P  V   F  T  D
2501  ATCCCTGTGG AGAACCTAGA GACAACCATG AGGTCCCCGG TGTTCACGGA
      TAGGGACACC TCTTGGATCT CTGTTGGTAC TCCAGGGGCC ACAAGTGCCT
```

```
    +2  N  S  S     P  P  V  V     P  Q  S     F  Q  V     A  H  L
2551    TAACTCCTCT CCACCAGTAG TGCCCCAGAG CTTCCAGGTG GCTCACCTCC
        ATTGAGGAGA GGTGGTCATC ACGGGGTCTC GAAGGTCCAC CGAGTGGAGG
```

```
    +2  H  A  P  T     G  S  G     K  S  T  K     V  P  A     A  Y  A
2601    ATGCTCCCAC AGGCAGCGGC AAAAGCACCA AGGTCCCGGC TGCATATGCA
        TACGAGGGTG TCCGTCGCCG TTTTCGTGGT TCCAGGGCCG ACGTATACGT
```

```
    +2  A  Q  G  ·Y     K  V  L     V  L  N     P  S  V  A     A  T  L·
2651    GCTCAGGGCT ATAAGGTGCT AGTACTCAAC CCCTCTGTTG CTGCAACACT
        CGAGTCCCGA TATTCCACGA TCATGAGTTG GGGAGACAAC GACGTTGTGA
```

```
    +2   G  F  G     A  Y  M     S  K  A  H     G  I  D     P  N  I
2701    GGGCTTTGGT GCTTACATGT CCAAGGCTCA TGGGATCGAT CCTAACATCA
        CCCGAAACCA CGAATGTACA GGTTCCGAGT ACCCTAGCTA GGATTGTAGT
```

```
    +2  R  T  G  V     R  T  I     T  T  G  S     P  I  T     Y  S  T
2751    GGACCGGGGT GAGAACAATT ACCACTGGCA GCCCCATCAC GTACTCCACC
        CCTGGCCCCA CTCTTGTTAA TGGTGACCGT CGGGGTAGTG CATGAGGTGG
```

```
    +2  Y  G  K  F     L  A  D     G  G  C     S  G  G  A     Y  D  I
2801    TACGGCAAGT TCCTTGCCGA CGGCGGGTGC TCGGGGGGCG CTTATGACAT
        ATGCCGTTCA AGGAACGGCT GCCGCCCACG AGCCCCCCGC GAATACTGTA
```

```
    +2  I  I  C     D  E  C  H     S  T  D     A  T  S     I  L  G
2851    AATAATTTGT GACGAGTGCC ACTCCACGGA TGCCACATCC ATCTTGGGCA
        TTATTAAACA CTGCTCACGG TGAGGTGCCT ACGGTGTAGG TAGAACCCGT
```

```
    +2  I  G  T  V     L  D  Q     A  E  T  A     G  A  R     L  V  V
2901    TTGGCACTGT CCTTGACCAA GCAGAGACTG CGGGGGCGAG ACTGGTTGTG
        AACCGTGACA GGAACTGGTT CGTCTCTGAC GCCCCCGCTC TGACCAACAC
```

```
    +2  L  A  T  A     T  P  P     G  S  V     T  V  P  H     P  N  I
2951    CTCGCCACCG CCACCCCTCC GGGCTCCGTC ACTGTGCCCC ATCCCAACAT
        GAGCGGTGGC GGTGGGGAGG CCCGAGGCAG TGACACGGGG TAGGGTTGTA
```

```
    +2  E  E  V     A  L  S  T     T  G  E     I  P  F     Y  G  K
3001    CGAGGAGGTT GCTCTGTCCA CCACCGGAGA GATCCCTTTT TACGGCAAGG
        GCTCCTCCAA CGAGACAGGT GGTGGCCTCT CTAGGGAAAA ATGCCGTTCC
```

```
    +2  A  I  P  L     E  V  I     K  G  G  R     H  L  I     F  C  H
3051    CTATCCCCCT CGAAGTAATC AAGGGGGGGA GACATCTCAT CTTCTGTCAT
        GATAGGGGGA GCTTCATTAG TTCCCCCCCT CTGTAGAGTA GAAGACAGTA
```

```
    +2  S  K  K  K     C  D  E     L  A  A     K  L  V  A     L  G  I
3101    TCAAAGAAGA AGTGCGACGA ACTCGCCGCA AAGCTGGTCG CATTGGGCAT
        AGTTTCTTCT TCACGCTGCT TGAGCGGCGT TTCGACCAGC GTAACCCGTA
```

```
    +2  N  A  V     A  Y  Y  R     G  L  D     V  S  V     I  P  T
3151    CAATGCCGTG GCCTACTACC GCGGTCTTGA CGTGTCCGTC ATCCCGACCA
        GTTACGGCAC CGGATGATGG CGCCAGAACT GCACAGGCAG TAGGGCTGGT
```

```
    +2  S  G  D  V     V  V  V     A  T  D  A     L  M  T     G  Y  T
3201    GCGGCGATGT TGTCGTCGTG GCAACCGATG CCCTCATGAC CGGCTATACC
        CGCCGCTACA ACAGCAGCAC CGTTGGCTAC GGGAGTACTG GCCGATATGG
```

```
      +2   G   D   F   D   S   V   I   D   C   N   T   C   V   T   Q   T   V
   3251   GGCGACTTCG ACTCGGTGAT AGACTGCAAT ACGTGTGTCA CCCAGACAGT
          CCGCTGAAGC TGAGCCACTA TCTGACGTTA TGCACACAGT GGGTCTGTCA
```

```
      +2   D   F   S   L   D   P   T   F   T   I   E   T   I   T   L   P
   3301   CGATTTCAGC CTTGACCCTA CCTTCACCAT TGAGACAATC ACGTCCCCC
          GCTAAAGTCG GAACTGGGAT GGAAGTGGTA ACTCTGTTAG TGCGAGGGGG
```

```
      +2   Q   D   A   V   S   R   T   Q   R   R   G   R   T   G   R   G   K
   3351   AAGATGCTGT CTCCCGCACT CAACGTCGGG GCAGGACTGG CAGGGGGAAG
          TTCTACGACA GAGGGCGTGA GTTGCAGCCC CGTCCTGACC GTCCCCCTTC
```

```
      +2   P   G   I   Y   R   F   V   A   P   G   E   R   P   S   G   M   F
   3401   CCAGGCATCT ACAGATTTGT GGCACCGGGG GAGCGCCCCT CCGGCATGTT
          GGTCCGTAGA TGTCTAAACA CCGTGGCCCC CTCGCGGGGA GGCCGTACAA
```

```
      +2   D   S   S   V   L   C   E   C   Y   D   A   G   C   A   W   Y
   3451   CGACTCGTCC GTCCTCTGTG AGTGCTATGA CGCAGGCTGT GCTTGGTATG
          GCTGAGCAGG CAGGAGACAC TCACGATACT GCGTCCGACA CGAACCATAC
```

```
      +2  E   L   T   P   A   E   T   T   V   R   L   R   A   Y   M   N   T
   3501   AGCTCACGCC CGCCGAGACT ACAGTTAGGC TACGAGCGTA CATGAACACC
          TCGAGTGCGG GCGGCTCTGA TGTCAATCCG ATGCTCGCAT GTACTTGTGG
```

```
      +2   P   G   L   P   V   C   Q   D   H   L   E   F   W   E   G   V   F
   3551   CCGGGGCTTC CCGTGTGCCA GGACCATCTT GAATTTTGGG AGGGCGTCTT
          GGCCCCGAAG GGCACACGGT CCTGGTAGAA CTTAAAACCC TCCCGCAGAA
```

```
      +2   T   G   L   T   H   I   D   A   H   F   L   S   Q   T   K   Q
           StuI
           ~~~~~~
   3601   TACAGGCCTC ACTCATATAG ATGCCCACTT TCTATCCCAG ACAAAGCAGA
          ATGTCCGGAG TGAGTATATC TACGGGTGAA AGATAGGGTC TGTTTCGTCT
```

```
      +2  S   G   E   N   L   P   Y   L   V   A   Y   Q   A   T   V   C   A
   3651   GTGGGGAGAA CCTTCCTTAC CTGGTAGCGT ACCAAGCCAC CGTGTGCGCT
          CACCCCTCTT GGAAGGAATG GACCATCGCA TGGTTCGGTG GCACACGCGA
```

```
      +2   R   A   Q   A   P   P   P   S   W   D   Q   M   W   K   C   L   I
   3701   AGGGCTCAAG CCCCTCCCCC ATCGTGGGAC CAGATGTGGA AGTGTTTGAT
          TCCCGAGTTC GGGGAGGGGG TAGCACCCTG GTCTACACCT TCACAAACTA
```

```
      +2   R   L   K   P   T   L   H   G   P   T   P   L   L   Y   R   L
   3751   TCGCCTCAAG CCCACCCTCC ATGGGCCAAC ACCCCTGCTA TACAGACTGG
          AGCGGAGTTC GGGTGGGAGG TACCCGGTTG TGGGGACGAT ATGTCTGACC
```

```
      +2  G   A   V   Q   N   E   I   T   L   T   H   P   V   T   K   Y   I
   3801   GCGCTGTTCA GAATGAAATC ACCCTGACGC ACCCAGTCAC CAAATACATC
          CGCGACAAGT CTTACTTTAG TGGGACTGCG TGGGTCAGTG GTTTATGTAG
```

```
      +2   M   T   C   M   S   A   D   L   E   V   V   T   S   T   W   V   L
   3851   ATGACATGCA TGTCGGCCGA CCTGGAGGTC GTCACGAGCA CCTGGGTGCT
          TACTGTACGT ACAGCCGGCT GGACCTCCAG CAGTGCTCGT GGACCCACGA
```

```
      +2   V   G   G   V   L   A   A   L   A   A   Y   C   L   S   T   G
   3901   CGTTGGCGGC GTCCTGGCTG CTTTGGCCGC GTATTGCCTG TCAACAGGCT
          GCAACCGCCG CAGGACCGAC GAAACCGGCG CATAACGGAC AGTTGTCCGA
```

```
      +2 C   V   V   I   V   G   R   V   V   L   S   G   K   P   A   I   I
      3951  GCGTGGTCAT AGTGGGCAGG GTCGTCTTGT CCGGGAAGCC GGCAATCATA
            CGCACCAGTA TCACCCGTCC CAGCAGAACA GGCCCTTCGG CCGTTAGTAT
```

```
      +2 P   D   R   E   V   L   Y   R   E   F   D   E   M   E   E   C
      4001  CCTGACAGGG AAGTCCTCTA CCGAGAGTTC GATGAGATGG AAGAGTGCTA
            GGACTGTCCC TTCAGGAGAT GGCTCTCAAG CTACTCTACC TTCTCACGAT
```

```
            BamHI   ·  MluI
            ~~~~~~     ~~~~~~~
      4051  GGATCCACTA CGCGTTAGAG CTCGCTGATC AGCCTCGACT GTGCCTTCTA
            CCTAGGTGAT GCGCAATCTC GAGCGACTAG TCGGAGCTGA CACGGAAGAT
```

```
      4101  GTTGCCAGCC ATCTGTTGTT TGCCCCTCCC CCGTGCCTTC CTTGACCCTG
            CAACGGTCGG TAGACAACAA ACGGGGAGGG GGCACGGAAG GAACTGGGAC
```

```
      4151  GAAGGTGCCA CTCCCACTGT CCTTTCCTAA TAAAATGAGG AAATTGCATC
            CTTCCACGGT GAGGGTGACA GGAAAGGATT ATTTTACTCC TTTAACGTAG
```

```
      4201  GCATTGTCTG AGTAGGTGTC ATTCTATTCT GGGGGGTGGG GTGGGGCAGG
            CGTAACAGAC TCATCCACAG TAAGATAAGA CCCCCCACCC CACCCCGTCC
```

```
      4251  ACAGCAAGGG GGAGGATTGG GAAGACAATA GCAGGCATGC TGGGGAGCTC
            TGTCGTTCCC CCTCCTAACC CTTCTGTTAT CGTCCGTACG ACCCCTCGAG
```

```
      4301  TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTCGTT CGGCTGCGGC
            AAGGCGAAGG AGCGAGTGAC TGAGCGACGC GAGCCAGCAA GCCGACGCCG
```

```
      4351  GAGCGGTATC AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA
            CTCGCCATAG TCGAGTGAGT TTCCGCCATT ATGCCAATAG GTGTCTTAGT
```

```
      4401  GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG
            CCCCTATTGC GTCCTTTCTT GTACACTCGT TTTCCGGTCG TTTTCCGGTC
```

```
      4451  GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC
            CTTGGCATTT TTCCGGCGCA ACGACCGCAA AAAGGTATCC GAGGCGGGGG
```

```
      4501  CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG
            GACTGCTCGT AGTGTTTTTA GCTGCGAGTT CAGTCTCCAC CGCTTTGGGC
```

```
      4551  ACAGGACTAT AAAGATACCA GGCGTTTCCC CCTGGAAGCT CCCTCGTGCG
            TGTCCTGATA TTTCTATGGT CCGCAAAGGG GGACCTTCGA GGGAGCACGC
```

```
      4601  CTCTCCTGTT CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC
            GAGAGGACAA GGCTGGGACG GCGAATGGCC TATGGACAGG CGGAAAGAGG
```

```
      4651  CTTCGGGAAG CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT
            GAAGCCCTTC GCACCGCGAA AGAGTTACGA GTGCGACATC CATAGAGTCA
```

```
      4701  TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT
            AGCCACATCC AGCAAGCGAG GTTCGACCCG ACACACGTGC TTGGGGGGCA
```

```
      4751  TCAGCCCGAC CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC
            AGTCGGGCTG GCGACGCGGA ATAGGCCATT GATAGCAGAA CTCAGGTTGG
```

```
      4801  CGGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT
            GCCATTCTGT GCTGAATAGC GGTGACCGTC GTCGGTGACC ATTGTCCTAA
```

```
4851  AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC
      TCGTCTCGCT CCATACATCC GCCACGATGT CTCAAGAACT TCACCACCGG

4901  TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA
      ATTGATGCCG ATGTGATCTT CCTGTCATAA ACCATAGACG CGAGACGACT

4951  AGCCAGTTAC CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA
      TCGGTCAATG GAAGCCTTTT TCTCAACCAT CGAGAACTAG GCCGTTTGTT

5001  ACCACCGCTG GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG
      TGGTGGCGAC CATCGCCACC AAAAAAACAA ACGTTCGTCG TCTAATGCGC

5051  CAGAAAAAAA GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG
      GTCTTTTTTT CCTAGAGTTC TTCTAGGAAA CTAGAAAAGA TGCCCCAGAC

5101  ACGCTCAGTG GAACGAAAAC TCACGTTAAG GGATTTTGGT CATGAGATTA
      TGCGAGTCAC CTTGCTTTTG AGTGCAATTC CCTAAAACCA GTACTCTAAT

5151  TCAAAAAGGA TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTTAA
      AGTTTTTCCT AGAAGTGGAT CTAGGAAAAT TTAATTTTTA CTTCAAAATT

5201  ATCAATCTAA AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT
      TAGTTAGATT TCATATATAC TCATTTGAAC CAGACTGTCA ATGGTTACGA

5251  TAATCAGTGA GGCACCTATC TCAGCGATCT GTCTATTTCG TTCATCCATA
      ATTAGTCACT CCGTGGATAG AGTCGCTAGA CAGATAAAGC AAGTAGGTAT

5301  GTTGCCTGAC TCCCCGTCGT GTAGATAACT ACGATACGGG AGGGCTTACC
      CAACGGACTG AGGGGCAGCA CATCTATTGA TGCTATGCCC TCCCGAATGG

5351  ATCTGGCCCC AGTGCTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC
      TAGACCGGGG TCACGACGTT ACTATGGCGC TCTGGGTGCG AGTGGCCGAG

5401  CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT
      GTCTAAATAG TCGTTATTTG GTCGGTCGGC CTTCCCGGCT CGCGTCTTCA

5451  GGTCCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAATT GTTGCCGGGA
      CCAGGACGTT GAAATAGGCG GAGGTAGGTC AGATAATTAA CAACGGCCCT

5501  AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTTGCCA
      TCGATCTCAT TCATCAAGCG GTCAATTATC AAACGCGTTG CAACAACGGT

5551  TTGCTACAGG CATCGTGGTG TCACGCTCGT CGTTTGGTAT GGCTTCATTC
      AACGATGTCC GTAGCACCAC AGTGCGAGCA GCAAACCATA CCGAAGTAAG

5601  AGCTCCGGTT CCCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG
      TCGAGGCCAA GGGTTGCTAG TTCCGCTCAA TGTACTAGGG GGTACAACAC

5651  CAAAAAAGCG GTTAGCTCCT TCGGTCCTCC GATCGTTGTC AGAAGTAAGT
      GTTTTTTCGC CAATCGAGGA AGCCAGGAGG CTAGCAACAG TCTTCATTCA

5701  TGGCCGCAGT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT
      ACCGGCGTCA CAATAGTGAG TACCAATACC GTCGTGACGT ATTAAGAGAA

5751  ACTGTCATGC CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC
      TGACAGTACG GTAGGCATTC TACGAAAAGA CACTGACCAC TCATGAGTTG
```

```
5801  CAAGTCATTC TGAGAATAGT GTATGCGGCG ACCGAGTTGC TCTTGCCCGG
      GTTCAGTAAG ACTCTTATCA CATACGCCGC TGGCTCAACG AGAACGGGCC

5851  CGTCAATACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC
      GCAGTTATGC CCTATTATGG CGCGGTGTAT CGTCTTGAAA TTTTCACGAG

5901  ATCATTGGAA AACGTTCTTC GGGGCGAAAA CTCTCAAGGA TCTTACCGCT
      TAGTAACCTT TTGCAAGAAG CCCCGCTTTT GAGAGTTCCT AGAATGGCGA

5951  GTTGAGATCC AGTTCGATGT AACCCACTCG TGCACCCAAC TGATCTTCAG
      CAACTCTAGG TCAAGCTACA TTGGGTGAGC ACGTGGGTTG ACTAGAAGTC

6001  CATCTTTTAC TTTCACCAGC GTTTCTGGGT GAGCAAAAAC AGGAAGGCAA
      GTAGAAAATG AAAGTGGTCG CAAAGACCCA CTCGTTTTTG TCCTTCCGTT

6051  AATGCCGCAA AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT
      TTACGGCGTT TTTTCCCTTA TTCCCGCTGT GCCTTTACAA CTTATGAGTA

6101  ACTCTTCCTT TTTCAATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA
      TGAGAAGGAA AAAGTTATAA TAACTTCGTA AATAGTCCCA ATAACAGAGT

6151  TGAGCGGATA CATATTTGAA TGTATTTAGA AAAATAAACA AATAGGGGTT
      ACTCGCCTAT GTATAAACTT ACATAAATCT TTTTATTTGT TTATCCCCAA

6201  CCGCGCACAT TTCCCCGAAA AGTGCCACCT GACGTCTAAG AAACCATTAT
      GGCGCGTGTA AAGGGGCTTT TCACGGTGGA CTGCAGATTC TTTGGTAATA

6251  TATCATGACA TTAACCTATA AAAATAGGCG TATCACGAGG CCCTTTCGTC
      ATAGTACTGT AATTGGATAT TTTTATCCGC ATAGTGCTCC GGGAAAGCAG
```

Diagram 1
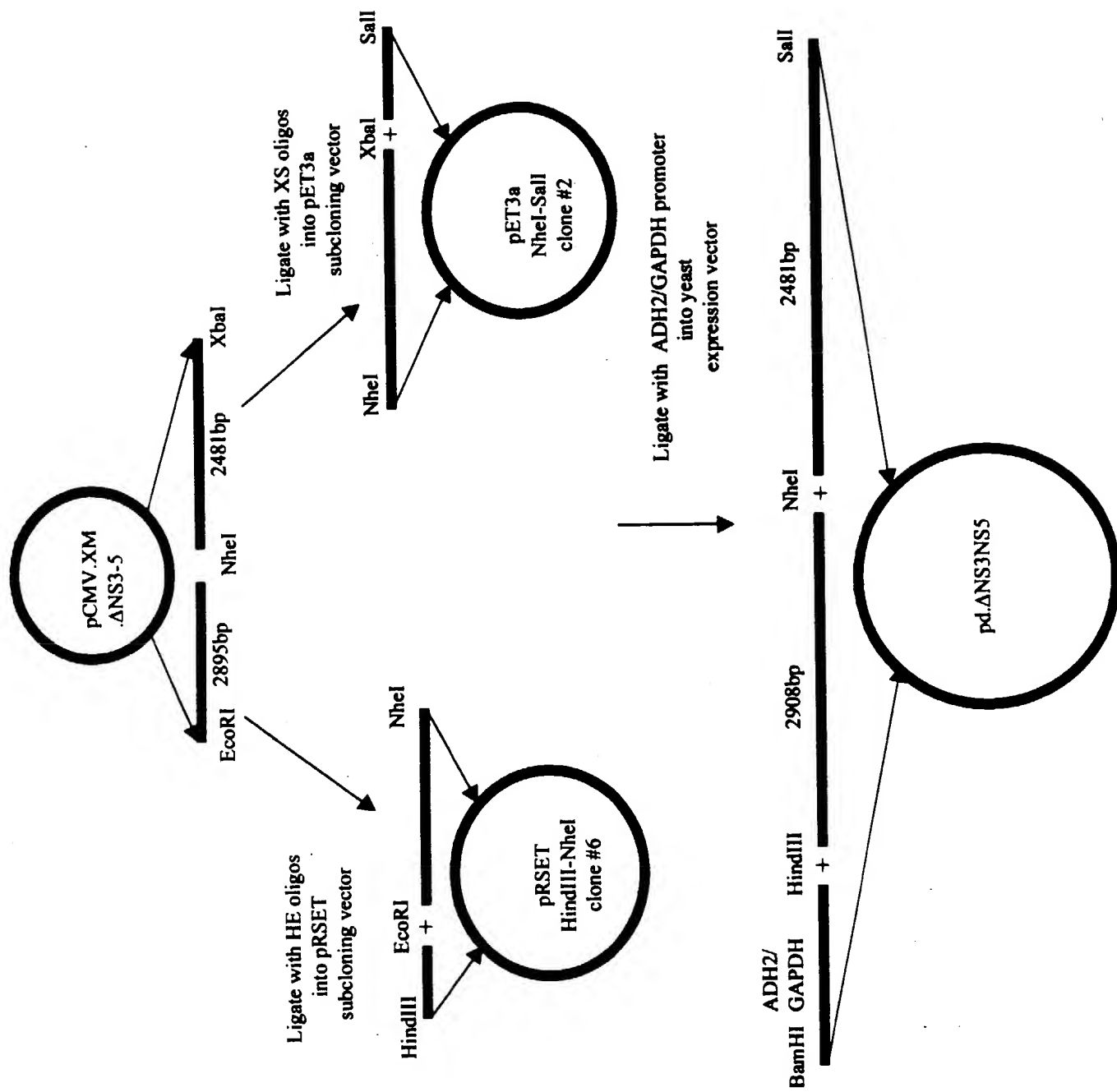
FIGURE 10

FIGURE 11 - Page 1

```
                            MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuVal
  2  AGCTTACAAAACAAATTCACCATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTA
     TCGAATGTTTTGTTTAAGTGGTACCGACGTATACGTCGAGTCCCGATATTCCACGATCAT
     ^              ^        ^                                 ^

  1 HIND3, 21 NCOI, 30 NDEI, 58 SCAI,


     LeuAsnProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGly
 62  CTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGG
     GAGTTGGGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCC


     IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyr
122  ATCGATCCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTAC
     TAGCTAGGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATG
     ^

122 CLAI,


     SerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIle
182  TCCACCTACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATA
     AGGTGGATGCCGTTCAAGGAACGGCTGCCGCCCACGAGCCCCCCGCGAATACTGTATTAT


     IleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeu
242  ATTTGTGACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTT
     TAAACACTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAA


     AspGlnAlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGly
302  GACCAAGCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGC
     CTGGTTCGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCG
                                       ^

309 ALWN1,


     SerValThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIle
362  TCCGTCACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATC
     AGGCAGTGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAG


     ProPheTyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePhe
422  CCTTTTTACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTC
     GGAAAAATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCCTCTGTAGAGTAGAAG


     CysHisSerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsn
482  TGTCATTCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAAT
     ACAGTAAGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTA


     AlaValAlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValVal
542  GCCGTGGCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTC
     CGGCACCGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAG
                                       ^        ^

556 SAC2, 566 DRD1,


     ValValAlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAsp
602  GTCGTGGCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGAC
     CAGCACCGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTG
                                       ^

621 BSPH1,


     CysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGlu
```

FIGURE 11 - Page 2

662 TGCAATACGTGTGTCACCCAGACAGTCGATTTCAGCCTTGACCCTACCTTCACCATTGAG
     ACGTTATGCACACAGTGGGTCTGTCAGCTAAAGTCGGAACTGGGATGGAAGTGGTAACTC

     ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArg
722 ACAATCACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGG
     TGTTAGTGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCC

     GlyLysProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAsp
782 GGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCCTCCGGCATGTTCGAC
     CCCTTCGGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTG
                              ^   ^

    822 BGLI, 839 DRD1,

     SerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAla
842 TCGTCCGTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCC
     AGCAGGCAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGG
                                                          ^

    887 SACI,

     GluThrThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAsp
902 GAGACTACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGAC
     CTCTGATGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTG
                                                  ^

    937 SMAI XMAI,

     HisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeu
962 CATCTTGAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTA
     GTAGAACTTAAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGAT
                                      ^

    991 STUI,

     SerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrVal
1022 TCCCAGACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTG
     AGGGTCTGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCAC
                                                               ^

    1075 DRA3,

     CysAlaArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArg
1082 TGCGCTAGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTCGC
     ACGCGATCCCGAGTTCGGGGAGGGGGTAGCACCCTGGTCTACACCTTCACAAACTAAGCG

     LeuLysProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsn
1142 CTCAAGCCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTCAGAAT
     GAGTTCGGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTA
                                  ^

    1156 NCOI,

     GluIleThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeu
1202 GAAATCACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTG
     CTTTAGTGGGACTGCGTGGGTCAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGAC
                                         ^   ^ ^   ^   ^

    1236 BSPH1, 1240 DRD1, 1243 AVA3, 1251 EAG1 XMA3, 1256 DRD1,

     GluValValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyr
1262 GAGGTCGTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTAT
     CTCCAGCAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGCGCATA

FIGURE 11 - Page 3

```
       CysLeuSerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAla
1322   TGCCTGTCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCA
       ACGGACAGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGT
                                                              ^
1375 NAEI,

       IleIleProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGln
1382   ATCATACCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAG
       TAGTATGGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTC
                  ^
1391 DRD1,

       HisLeuProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeu
1442   CACTTACCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTC
       GTGAATGGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAG

       GlyLeuLeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsn
1502   GGCCTCCTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAAC
       CCGGAGGACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTG
            ^        ^
1508 PSTI,  1513 TTH3I,

       TrpGlnLysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGln
1562   TGGCAAAAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACTTCATCAGTGGGATACAA
       ACCGTTTTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTT
                  ^                        ^
1571 XHOI,  1592 NDEI,

       TyrLeuAlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPhe
1622   TACTTGGCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTT
       ATGAACCGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACTACCGAAAA
                                   ^
1649 BSTE2,

       ThrAlaAlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGly
1682   ACAGCTGCTGTCACCAGCCCACTAACCACTAGCCAAACCCTCCTCTTCAACATATTGGGG
       TGTCGACGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCC
            ^
1683 ALWN1 PVU2,

       GlyTrpValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGly
1742   GGGTGGGTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGC
       CCCACCCACCGACGGGTCGAGCGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCG
                                                                 ^
1800 ESP1,

       LeuAlaGlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAla
1802   TTAGCTGGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGCA
       AATCGACCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGT
            ^
1808 KAS1 NARI,

       GlyTyrGlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluVal
1862   GGGTATGGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTC
       CCCATACCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAG
                             ^                        ^
```

FIGURE 11 - Page 4

1884 SACI, 1905 BSPH1,

```
    ProSerThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuVal
1922 CCCTCCACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTA
    GGGAGGTGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCAT
                           ^
```

1934 TTH3I,

```
    ValGlyValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaVal
1982 GTCGGCGTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTG
    CAGCCGCACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCAC
                       ^                        ^
```

2010 NAEI, 2023 SMAI XMAI,

```
    GlnTrpMetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHis
2042 CAGTGGATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCCACGCAC
    GTCACCTACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTG
                           ^                            ^
```

2073 SMAI XMAI, 2099 DRA3,

```
    TyrValProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrVal
2102 TACGTGCCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTA
    ATGCACGGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACAT
                       ^
```

2121 PVU2,

```
    ThrGlnLeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSer
2162 ACCCAGCTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCC
    TGGGTCGAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGG
         ^     ^
```

2165 ALWN1, 2170 MST2,

```
    GlySerTrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThr
2222 GGTTCCTGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACC
    CCAAGGACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACTCGCTGAAATTCTGG
         ^
```

2226 ECON1,

```
    TrpLeuLysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArg
2282 TGGCTAAAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGC
    ACCGATTTTCGATTCGAGTACGGTGTCGACGGACCCTAGGGGAAACACAGGACGGTCGCG
                 ^                  ^              ^
```

2291 ESP1, 2306 PVU2, 2316 BAMHI,

```
    GlyTyrLysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAla
2342 GGGTATAAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCT
    CCCATATTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGA
```

```
    GluIleThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArg
2402 GAGATCACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGG
    CTCTAGTGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCC
                           ^                        ^        ^^
```

2431 BSAB1, 2447 AVR2, 2454 SSE83871, 2455 PSTI,

```
    AsnMetTrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeu
2462 AACATGTGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTT
    TTGTACACCTCACCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAA
```

FIGURE 11 - Page 5

2486 ASE1, 2503 APAI,

```
      ProAlaProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIle
2522  CCTGCGCCGAACTACACGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATA
      GGACGCGGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTAT
```

2559 PSTI,

```
      ArgGlnValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysPro
2582  AGGCAGGTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCG
      TCCGTCCACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGC
```

2600 DRA3,

```
      CysGlnValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPhe
2642  TGCCAGGTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTT
      ACGGTCCAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAA
```

```
      AlaProProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGlu
2702  GCGCCCCCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTCAGAGTAGGACTCCACGAA
      CGCGGGGGGACGTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTT
```

```
      TyrProValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSer
2762  TACCCGGTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCC
      ATGGGCCATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACTGCAGG
```

2763 HGIE2, 2815 AAT2,

```
      MetLeuThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGly
2822  ATGCTCACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGA
      TACGAGTGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCCT
```

2856 EAG1 XMA3,

```
      SerProProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAla
2882  TCACCCCCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCA
      AGTGGGGGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGT
```

2895 BALI, 2909 NHEI,

```
      ThrCysThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrp
2942  ACTTGCACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGG
      TGAACGTGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACC
```

2972 ESP1, 2975 SACI,

```
      ArgGlnGluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeu
3002  AGGCAGGAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTG
      TCCGTCCTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTTCACCACTAAGAC
```

```
      AspSerPheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGlu
3062  GACTCCTTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAA
      CTGAGGAAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTT
```

3102 BGL2,

FIGURE 11 - Page 6

```
        IleLeuArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyr
3122    ATCCTGCGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTAT
        TAGGACGCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATA
                                   ^                          ^
3149 ALWN1, 3170 EAG1 XMA3,


        AsnProProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGly
3182    AACCCCCCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGC
        TTGGGGGGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCG
                                                  ^            ^
3223 HGIE2, 3235 NCOI,


        CysProLeuProProProLysSerProProValProProProArgLysLysArgThrVal
3242    TGCCCGCTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTG
        ACGGGCGAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCAC


        ValLeuThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGly
3302    GTCCTCACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGC
        CAGGAGTGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCG
                                                  ^            ^
3338 SACI, 3352 HIND3,


        SerSerSerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaPro
3362    AGCTCCTCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCCT
        TCGAGGAGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGA


        SerGlyCysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGly
3422    TCTGGCTGCCCCCCCGACTCCGACGCTGAGTCCTATTCCTCCATGCCCCCCCTGGAGGGG
        AGACCGACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCC
                                   ^
3443 EAM11051,


        GluProGlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsn
3482    GAGCCTGGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAAC
        CTCGGACCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTG
                   ^^ ^
3490 BAMHI, 3491 BSAB1, 3493 BSPE1,


        AlaGluAspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrPro
3542    GCGGAGGATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCG
        CGCCTCCTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGC
                                                               ^
3595 DRA3,


        CysAlaAlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHis
3602    TGCGCCGCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCAC
        ACGCGGCGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTG
           ^            ^                                         ^
3606 SAC2, 3617 ALWN1, 3661 PFLM1,


        HisAsnLeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThr
3662    CACAATTTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACA
        GTGTTAAACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGT
                                   ^
3687 DRA3,


        PheAspArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAla
```

FIGURE 11 - Page 7

3722 TTTGACAGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCA
     AAACTGTCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCGT

     AlaAlaSerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrPro
3782 GCGGCGTCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCC
     CGCCGCAGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGG
                                                             ^

3822 HIND3,

     ProHisSerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArg
3842 CCACACTCAGCCAAATCCAAGTTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGA
     GGTGTGAGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCT
                                     ^   ^

3881 AAT2, 3896 BGLI,

     LysAlaValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrPro
3902 AAGGCCGTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCA
     TTCCGGCATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGT

     IleAspThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGly
3962 ATAGACACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGT
     TATCTGTGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCCA

     ArgLysProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMet
4022 CGTAAGCCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATG
     GCATTCGGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTAC

     AlaLeuTyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPhe
4082 GCTTTGTACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTC
     CGAAACATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAG

     GlnTyrSerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThr
4142 CAATACTCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACC
     GTTATGAGTGGTCCTGTCGCCCAACTTAAGGAGCACGTTCGCACCTTCAGGTTCTTTTGG
                                   ^

4166 ECORI,

     ProMetGlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIle
4202 CCAATGGGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATC
     GGTTACCCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAG
                                         ^   ^

4235 DRD1, 4242 ALWN1,

     ArgThrGluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIle
4262 CGTACGGAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCGTGGCCATC
     GCATGCCTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCGGGCGCACCGGTAG
                                                   ^   ^

4307 BGLI, 4314 BALI,

     LysSerLeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsn
4322 AAGTCCCTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAAC
     TTCAGGGAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTTCCCCCCTCTTG
                                            ^

4351 APAI,

     CysGlyTyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeu
4382 TGCGGCTATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTC

FIGURE 11 - Page 8

ACGCCGATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAG

```
          ThrCysTyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMet
     4442 ACTTGCTACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATG
          TGAACGATGTAGTTCCGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTAC
                                 ^
```

4458 SMAI XMAI,

```
          LeuValCysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAla
     4502 CTCGTGTGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCG
          GAGCACACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGC
                                 ^   ^
```

4514 DRD1, 4517 TTH3I,

```
          AlaSerLeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspPro
     4562 GCGAGCCTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCCTGGGGACCCC
          CGCTCGGACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGGACCCCTGGGG
```

```
          ProGlnProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAla
     4622 CCACAACCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCC
          GGTGTTGGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGG
                                       ^
```

4643 SACI,

```
          HisAspGlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAla
     4682 CACGACGGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCG
          GTGCTGCCGCGACCTTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGC
                                                             ^
```

4737 NRUI,

```
          ArgAlaAlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIle
     4742 AGAGCTGCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATC
          TCTCGACGCACCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAG
```

```
          MetPheAlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeu
     4802 ATGTTTGCCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTT
          TACAAACGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAA
                                 ^^
```

4812 PFLM1, 4813 DRA3,

```
          IleAlaArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSer
     4862 ATAGCCAGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCC
          TATCGGTCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGG
                                                   ^
```

4899 BGL2,

```
          IleGluProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSer
     4922 ATAGAACCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCA
          TATCTTGGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGT
                                                   ^
```

4960 NCOI,

```
          LeuHisSerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGly
     4982 CTCCACAGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACTTGGG
          GAGGTGTCAATGAGAGGTCCACTTTAGTTATCCCACCGGCGTACGGAGTCTTTTGAACCC
                                                   ^                     ^
```

5021 SPHI, 5041 KPNI,

FIGURE 11 - Page 9

```
       ValProProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAla
5042   GTACCGCCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCC
       CATGGCGGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGG
                                    ^                            ^

5070   APAI, 5097 BALI,

       ArgGlyGlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLys
5102   AGAGGAGGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAG
       TCTCCTCCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTC
                         ^

5119   NDEI,

       LeuLysLeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAla
5162   CTCAAACTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCT
       GAGTTTGAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGA
                             ^^       ^    ^

5180   NOTI, 5181 EAG1 XMA3, 5188 BALI, 5192 PVU2,

       GlyTyrSerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrp
5222   GGCTACAGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCCGCTGGATCTGG
       CCGATGTCGCCCCCTCTGTAAATAGTGTCGCACAGAGTACGGGCCGGGGCGACCTAGACC
                        ^

5246   DRA3,

       PheCysLeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP
5282   TTTTGCCTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGATGAAGG
       AAAACGGATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTACTTCC
                             ^                            ^

5301   PSTI, 5331 HGIE2,


5342   TTGGGGTAAACACTCCGGCCTAAAAAAAAAAAAAAAATCTAGAACCCGAGTCGAC
       AACCCCATTTGTGAGGCCGGATTTTTTTTTTTTTTTTAGATCTTGGGCTCAGCTG
                                            ^           ^

5378   XBAI, 5390 SALI,
```

FIGURE 12

FIGURE 13

FIGURE 14 - Page 1

```
                    MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn
  2  AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
     TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG
     ^              ^                                    ^
  1 HIND3, 24 NDEI, 52 SCAI,


                    ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp
 62  CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT
     GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA
                                                         ^
 116 CLAI,


                    ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr
122  CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC
     GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG


                    TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys
182  TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT
     ATGCCGTTCAAGGAACGGCTGCCGCCCACGAGCCCCCCGCGAATACTGTATTATTAAACA


                    AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln
242  GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA
     CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAACTGGTT


                    AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal
302  GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC
     CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG
     ^
 303 ALWN1,


                    ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe
362  ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT
     TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA


                    TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
422  TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTCAT
     ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCCTCTGTAGAGTAGAAGACAGTA


                    SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
482  TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG
     AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC


                    AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
542  GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTCGTCGTG
     CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC
                 ^          ^
 550 SAC2, 560 DRD1,


                    AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn
602  GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT
     CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA
                    ^
 615 BSPH1,


     ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle
```

FIGURE 14 - Page

```
662  ACGTGTGTCACCCAGACAGTCGATTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC
     TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAACTGGGATGGAAGTGGTAACTCTGTTAG


     ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
722  ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG
     TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC


     ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
782  CCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCCTCCGGCATGTTCGACTCGTCC
     GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG
                                         ^                    ^

816 BGLI, 833 DRD1,


     ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
842  GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT
     CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA
                                              ^

881 SACI,


     ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
902  ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT
     TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAA
                                       ^

931 SMAI XMAI,


     GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
962  GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG
     CTTAAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC
                                  ^

985 STUI,


     ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
1022 ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT
     TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA
                                                        ^

1069 DRA3,


     ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
1082 AGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTCGCCTCAAG
     TCCCGAGTTCGGGGAGGGGGTAGCACCCTGGTCTACACCTTCACAAACTAAGCGGAGTTC


     ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
1142 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTCAGAATGAAATC
     GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG
                  ^

1150 NCOI,


     ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC
     TGGGACTGCGTGGGTCAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG
                            ^    ^  ^              ^      ^

1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,


     ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG
     CAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGCGCATAACGGAC
```

FIGURE 14 - Page

```
          SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
     1322 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA
          AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT
                                                               ^
     1369 NAEI,


          ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
     1382 CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA
          GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT
             ^
     1385 DRD1,


          ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
     1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC
          GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG


          LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
     1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA
          GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
          ^         ^
     1502 PSTI, 1507 TTH3I,


          LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
     1562 AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACTTCATCAGTGGGATACAATACTTG
          TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC
             ^                            ^
     1565 XHOI, 1586 NDEI,


          AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
     1622 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
          CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCGA
                       ^                                          ^
     1643 BSTE2, 1677 ALWN1 PVU2,


          AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
     1682 GCTGTCACCAGCCCACTAACCACTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG
          CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCCACC


          ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
     1742 GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
          CACCGACGGGTCGAGCGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA
                                                          ^
     1794 ESP1,


          GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
     1802 GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT
          CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA
          ^
     1802 KAS1 NARI,


          GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
     1862 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC
          CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
                       ^                            ^
     1878 SACI, 1899 BSPH1,
```

FIGURE 14 - Page 4

```
      ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
1922  ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC
      TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
                                 ^
      1928 TTH3I,


      ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
1982  GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG
      CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC
                                 ^                  ^
      2004 NAEI, 2017 SMAI XMAI,


      MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
2042  ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCCACGCACTACGTG
      TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC
                                 ^                           ^
      2067 SMAI XMAI, 2093 DRA3,


      ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
2102  CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG
      GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC
                             ^                                     ^
      2115 PVU2, 2159 ALWN1,


      LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
2162  CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC
      GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
           ^                                                   ^
      2164 MST2, 2220 ECON1,


      TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
2222  TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA
      ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACTCGCTGAAATTCTGGACCGAT


      LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
2292  AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT
      TTTCGATTCGAGTACGGTGTCGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCATA
            ^                    ^                ^
      2285 ESP1, 2300 PVU2, 2310 BAMHI,


      LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
2342  AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
      TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG


      ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
2402  ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG
      TGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGTAC
                                        ^              ^      ^^
      2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,


      TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
2462  TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCCTGCG
      ACCTCACCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC
                            ^                    ^
      2480 ASE1, 2497 APAI,


      ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln
```

FIGURE 14 - Page 5

2522   CCGAACTACACGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG
        GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC

2553 PSTI,

       ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
2582   GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG
        CACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC

2594 DRA3,

       ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
2642   GTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC
        CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG

       ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
2702   CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTCAGAGTAGGACTCCACGAATACCCG
        GGGACGTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC

2757 HGIE2,

       ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
2762   GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC
        CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACTGCAGGTACGAG

2809 AAT2,

       ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
2822   ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGATCACCC
        TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCCTAGTGGG

2850 EAG1 XMA3,

       ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys
2832   CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
        GGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG

2889 BALI, 2903 NHEI,

       ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
2942   ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
        TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC

2966 ESP1, 2969 SACI,

       GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
3002   GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC
        CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTTCACCACTAAGACCTGAGG

       PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
3062   TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG
        AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC

3096 BGL2,

       ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
3122   CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC

FIGURE 14 - Page 6

GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG

3143 ALWN1, 3164 EAG1 XMA3,


```
      ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
3182  CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG
      GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC
```

3217 HGIE2, 3229 NCOI,


```
      LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
3242  CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC
      GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG
```


```
      ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
3302  ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC
      TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG
```

3332 SACI, 3346 HIND3,


```
      SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
3362  TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCCTTCTGGC
      AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG
```


```
      CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
3422  TGCCCCCCCGACTCCGACGCTGAGTCCTATTCCTCCATGCCCCCCCTGGAGGGGGAGCCT
      ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA
```

3437 EAM11051,


```
      GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
3482  GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG
      CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC
```

3484 BAMHI, 3485 BSAB1, 3487 BSPE1,


```
      AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
3542  GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC
      CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG
```

3589 DRA3, 3600 SAC2,


```
      AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
3602  GCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT
      CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTGGTGTTA
```

3611 ALWN1, 3655 PFLM1,


```
      LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
3662  TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC
      AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAACTG
```

3681 DRA3,


```
      ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
3722  AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG
      TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCGTCGCCGC
```

FIGURE 14 - Page 7

```
        SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
3782    TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCCACAC
        AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG
                                     ^

3816 HIND3,

        SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
3842    TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC
        AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG
                                ^                          ^

3875 AAT2, 3890 BGLI,

        ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
3902    GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC
        CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

        ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
3962    ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG
        TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCCAGCATTC

        ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
4022    CCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG
        GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

        TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
4082    TACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC
        ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

        SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
4142    TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG
        AGTGGTCCTGTCGCCCAACTTAAGGAGCACGTTCGCACCTTCAGGTTCTTTTGGGGTTAC
                           ^

4160 ECORI,

        GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
4202    GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
        CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC
                                    ^            ^

4229 DRD1, 4236 ALWN1,

        GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
4262    GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCGTGGCCATCAAGTCC
        CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCGGGCGCACCGGTAGTTCAGG
                                                ^           ^

4301 BGLI, 4308 BALI,

        LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
4322    CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC
        GAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTTCCCCCCTCTTGACGCCG
                           ^

4345 APAI,

        TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys
4382    TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC
        ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG
```

FIGURE 14 - Page 8

```
        TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
4442    TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG
        ATGTAGTTCCGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC
                  ^
        4452 SMAI XMAI,


        CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
4502    TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC
        ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG
                  ^  ^
        4508 DRD1, 4511 TTH3I,


        LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
4562    CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCCTGGGGACCCCCCACAA
        GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGGACCCCTGGGGGGTGTT


        ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp
4622    CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC
        GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG
                          ^
        4637 SACI,


        GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
4682    GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT
        CCGCGACCTTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA
                                                            ^
        4731 NRUI,


        AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
4742    GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT
        CGCACCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA


        AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla
4802    GCCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC
        CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG
                  ^^
        4806 PFLM1, 4807 DRA3,


        ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu
4862    AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGCCTGCTACTCCATAGAA
        TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT
                                      ^
        4893 BGL2,


        ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
4922    CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC
        GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG
                                              ^
        4954 NCOI,


        SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
4982    AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACTTGGGGTACCG
        TCAATGAGAGGTCCACTTTAGTTATCCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC
                                          ^                     ^
        5015 SPHI, 5035 KPNI,


        ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly
```

FIGURE 14 - Page 9

5042   CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
          GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT

5064 APAI, 5091 BALI,

        GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA
          CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT

5113 NDEI,

        LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC
          GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

        SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCCGCTGGATCTGGTTTTGC
          TCGCCCCCTCTGTAAATAGTGTCGCACAGAGTACGGGCCGGGGCGACCTAGACCAAAACG

5240 DRA3,

        LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgOP
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGATGAATAGTCGAC
          GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTACTTATCAGCTG

5295 PSTI, 5336 SALI,
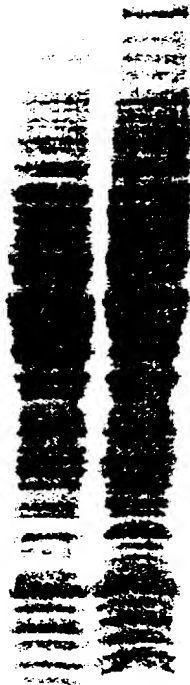
FIGURE 15

SCANNED, # _14_

FIGURE 16 - Page 1

FIGURE 16 - Pa

NotI

NotI

pSP72
HindIII-SalI
clone #6

121bp

PstI

PstI

pd.ΔNS3NS5.PJ
core121RT

403bp

SalI

Ligate fragments into pd.NS3NS5.PJ
NotI-SalI cloning vector.

NotI

PstI

pd.ΔNS3NS5.PJ
core173RT

559bp

SalI

FIGURE 17 - Page 1

```
                    MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn
  2  AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
     TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG
     ^                          ^                              ^
  1 HIND3, 24 NDEI, 52 SCAI,

     ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp
 62  CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT
     GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA
                                                          ^
     116 CLAI,

     ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr
 122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC
     GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG

     TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys
 182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT
     ATGCCGTTCAAGGAACGGCTGCCGCCCACGAGCCCCCCGCGAATACTGTATTATTAAACA

     AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln
 242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA
     CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAACTGGTT

     AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal
 302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC
     CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG
        ^
     303 ALWN1,

     ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe
 362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT
     TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA

     TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
 422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTCAT
     ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCCTCTGTAGAGTAGAAGACAGTA
```

FIGURE 17 - Page 2

```
         SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
    482  TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG
         AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC

         AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
    542  GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTCGTCGTG
         CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC
                       ^                ^
    550 SAC2, 560 DRD1,

         AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn
    602  GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT
         CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA
                             ^
    615 BSPH1,

         ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle
    662  ACGTGTGTCACCCAGACAGTCGATTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC
         TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAACTGGGATGGAAGTGGTAACTCTGTTAG

         ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
    722  ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGGCAGGACTGGCAGGGGGAAG
         TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC

         ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
    782  CCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCCTCCGGCATGTTCGACTCGTCC
         GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG
                                        ^                ^
    816 BGLI, 833 DRD1,

         ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
    842  GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT
         CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA
                                        ^
    881 SACI,

         ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
    902  ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT
         TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAA
                                    ^
    931 SMAI XMAI,

         GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
    962  GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG
         CTTAAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC
                                          ^
    985 STUI,

         ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
   1022  ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT
         TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA
                                                      ^
   1069 DRA3,

         ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
   1082  AGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTCGCCTCAAG
```

FIGURE 17 - Page 3

TCCCGAGTTCGGGGAGGGGGTAGCACCCTGGTCTACACCTTCACAAACTAAGCGGAGTTC

```
     ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
1142 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTCAGAATGAAATC
     GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG
                   ^
1150 NCOI,

     ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC
     TGGGACTGCGTGGGTCAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG
                         ^      ^   ^              ^        ^
1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,

     ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG
     CAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGCGCATAACGGAC

     SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
1322 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA
     AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT
                                                          ^
1369 NAEI,

     ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
1382 CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA
     GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT
               ^
1385 DRD1,

     ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC
     GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

     LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA
     GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
     ^        ^
1502 PSTI, 1507 TTH3I,

     LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
1562 AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACTTCATCAGTGGGATACAATACTTG
     TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC
            ^                             ^
1565 XHOI, 1586 NDEI,

     AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
1622 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
     CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCGA
                       ^                                    ^
1643 BSTE2, 1677 ALWN1 PVU2,

     AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
1682 GCTGTCACCAGCCCACTAACCACTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG
     CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCCACC
```

FIGURE 17 - Page

```
       ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
1742   GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
       CACCGACGGGTCGAGCGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA
                                                                 ^
       1794 ESP1,


       GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
1802   GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT
       CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA
       ^
       1802 KAS1 NARI,


       GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
1862   GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC
       CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
                          ^                          ^
       1878 SACI, 1899 BSPH1,


       ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
1922   ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC
       TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
                ^
       1928 TTH3I,


       ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
1982   GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG
       CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCGTCACGTCACC
                         ^                  ^
       2004 NAEI, 2017 SMAI XMAI,


       MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
2042   ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCCACGCACTACGTG
       TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC
                         ^                            ^
       2067 SMAI XMAI, 2093 DRA3,


       ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
2102   CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG
       GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC
                     ^                                           ^
       2115 PVU2, 2159 ALWN1,


       LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
2162   CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC
       GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
       ^                                                         ^
       2164 MST2, 2220 ECON1,


       TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
2222   TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA
       ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACTCGCTGAAATTCTGGACCGAT


       LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
2282   AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT
       TTTCGATTCGAGTACGGTGTCGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCATA
       ^                  ^               ^
       2285 ESP1, 2300 PVU2, 2310 BAMHI,
```

FIGURE 17 - Page 5

```
     LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
2342 AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
     TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG
                          ^                    ^      ^^

     2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

     TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
2462 TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCCTGCG
     ACCTCACCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC
                       ^                       ^

     2480 ASE1, 2497 APAI,

     ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln
2522 CCGAACTACACGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG
     GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC
                                          ^

     2553 PSTI,

     ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
2582 GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG
     CACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC
                       ^

     2594 DRA3,

     ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
2642 GTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC
     CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG

     ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
2702 CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTCAGAGTAGGACTCCACGAATACCCG
     GGGACGTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC
                                                            ^

     2757 HGIE2,

     ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
2762 GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC
     CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACTGCAGGTACGAG
                                                         ^

     2809 AAT2,

     ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
2822 ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGATCACCC
     TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCCTAGTGGG
                      ^

     2850 EAG1 XMA3,

     ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys
2882 CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
     GGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG
              ^                 ^

     2889 BALI, 2903 NHEI,
```

FIGURE 17 - Page 6

```
      ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
2942  ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
      TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC
                                     ^   ^
      2966 ESP1, 2969 SACI,


      GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
3002  GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC
      CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTTCACCACTAAGACCTGAGG


      PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
3062  TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG
      AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC
                                          ^
      3096 BGL2,


      ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
3122  CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC
      GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG
                      ^                            ^
      3143 ALWN1, 3164 EAG1 XMA3,


      ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
3182  CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG
      GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC
                                                 ^          ^
      3217 HGIE2, 3229 NCOI,


      LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
3242  CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC
      GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG


      ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
3302  ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC
      TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG
                                        ^          ^
      3332 SACI, 3346 HIND3,


      SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
3362  TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCCTTCTGGC
      AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG


      CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
3422  TGCCCCCCCGACTCCGACGCTGAGTCCTATTCCTCCATGCCCCCCCCTGGAGGGGGAGCCT
      ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA
                    ^
      3437 EAM11051,


      GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
3482  GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG
      CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC
      ^^  ^
      3484 BAMHI, 3485 BSAB1, 3487 BSPE1,


      AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
3542  GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC
      CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG
```

FIGURE 17 - Page 7

3589 DRA3, 3600 SAC2,

```
      AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
3602  GCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT
      CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTGGTGTTA
                   ^                                            ^
```

3611 ALWN1, 3655 PFLM1,

```
      LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
3662  TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC
      AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAACTG
                          ^
```

3681 DRA3,

```
      ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
3722  AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG
      TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCGTCGCCGC
```

```
      SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
3782  TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCCACAC
      AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG
                                                  ^
```

3816 HIND3,

```
      SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
3842  TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC
      AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG
                                    ^                    ^
```

3875 AAT2, 3890 BGLI,

```
      ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
3902  GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC
      CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG
```

```
      ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
3962  ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG
      TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCCAGCATTC
```

```
      ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
4022  CCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG
      GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC
```

```
      TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
4082  TACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC
      ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG
```

```
      SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
4142  TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG
      AGTGGTCCTGTCGCCCAACTTAAGGAGCACGTTCGCACCTTCAGGTTCTTTTGGGGTTAC
                                 ^
```

4160 ECORI,

```
      GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
4202  GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
      CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC
                              ^          ^
```

FIGURE 17 - Page 8

4229 DRD1, 4236 ALWN1,

```
     GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
4262 GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCGTGGCCATCAAGTCC
     CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCGGGCGCACCGGTAGTTCAGG
                                          ^         ^
```

4301 BGLI, 4308 BALI,

```
     LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
4322 CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC
     GAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTTCCCCCCTCTTGACGCCG
                    ^
```

4345 APAI,

```
     TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys
4382 TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC
     ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG
```

```
     TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
4442 TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG
     ATGTAGTTCCGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC
                      ^
```

4452 SMAI XMAI,

```
     CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
4502 TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC
     ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG
                  ^   ^
```

4508 DRD1, 4511 TTH3I,

```
     LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
4562 CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCCTGGGGACCCCCCACAA
     GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGGGACCCCTGGGGGGTGTT
```

```
     ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp
4622 CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC
     GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG
                          ^
```

4637 SACI,

```
     GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
4682 GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT
     CCGCGACCTTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA
                                                     ^
```

4731 NRUI,

```
     AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
4742 GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT
     CGCACCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA
```

```
     AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla
4802 GCCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC
     CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG
                  ^^
```

4806 PFLM1, 4807 DRA3,

```
     ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu
```

FIGURE 17 - Page 9

4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA
TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT
                                        ^

4893 BGL2,

     ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC
     GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG
                                        ^

4954 NCOI,

     SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACTTGGGGTACCG
     TCAATGAGAGGTCCACTTTAGTTATCCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC
                                        ^   ^

5015 SPHI, 5035 KPNI,

     ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly
5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
     GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT
                              ^   ^

5064 APAI, 5091 BALI,

     GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA
     CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT
                    ^

5113 NDEI,

     LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC
     GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG
                      ^^   ^   ^

5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

     SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCCGCTGGATCTGGTTTTGC
     TCGCCCCCTCTGTAAATAGTGTCGCACAGAGTACGGGCCGGGGCGACCTAGACCAAAACG
                    ^

5240 DRA3,

     LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGAATGAGCACGAAT
     GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA
                  ^

5295 PSTI,

     ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe
5342 CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCGGCCGCAGGACGTCAAGTTC
     GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGGCGTCCTGCAGTTCAAG
                                        ^^   ^   ^

5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

     ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu
5402 CCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTTGCCGCGCAGGGGCCCTAGATTG
     GGCCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC
                                        ^

FIGURE 17 - Page 10

5449 APAI,

```
        GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro
5462    GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT
        CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA
                ^               ^                        ^            ^
```

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

```
        IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
5522    ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC
        TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGG
                                    ^                ^        ^    ^
```

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

```
        LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
5582    CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG
        GAGATACCGTTACTCCCGACGCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC
```

```
        ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysOC AM
5642    CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTCGCGCAATTTGGGTAAGTAATAGTCG
        GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCATTATCAGC
                ^                                                    ^
```

5650 APAI, 5698 SALI,

5702   AC
        TG

FIGURE 18 - Page 1

```
                        MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn
      2  AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
         TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG
         ^                        ^                              ^

      1 HIND3, 24 NDEI, 52 SCAI,


         ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp
     62  CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT
         GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA
                                                                  ^

      116 CLAI,


         ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr
     122 CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC
         GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG


         TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys
     182 TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT
         ATGCCGTTCAAGGAACGGCTGCCGCCCACGAGCCCCCGCGAATACTGTATTATTAAACA


         AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln
     242 GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA
         CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAACTGGTT


         AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal
     302 GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC
         CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG
                                                              ^

      303 ALWN1,


         ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe
     362 ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT
         TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA


         TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
     422 TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTCAT
         ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCCTCTGTAGAGTAGAAGACAGTA


         SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
     482 TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG
         AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC


         AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
     542 GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTCGTCGTG
         CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC
                                ^          ^

      550 SAC2, 560 DRD1,


         AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn
     602 GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT
         CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA
                                ^

      615 BSPH1,
```

FIGURE 18 - Page 2

```
         ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle
     662 ACGTGTGTCACCCAGACAGTCGATTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC
         TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAACTGGGATGGAAGTGGTAACTCTGTTAG


         ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
     722 ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG
         TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC


         ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
     782 CCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCCTCCGGCATGTTCGACTCGTCC
         GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG
                                                ^                  ^

         816 BGLI, 833 DRD1,


         ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
     842 GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT
         CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA
                                                      ^

         881 SACI,


         ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
     902 ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT
         TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAA
                                         ^

         931 SMAI XMAI,


         GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
     962 GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG
         CTTAAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC
                                     ^

         985 STUI,


         ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
    1022 ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT
         TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA
                                                          ^

         1069 DRA3,


         ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
    1082 AGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTCGCCTCAAG
         TCCCGAGTTCGGGGAGGGGGTAGCACCCTGGTCTACACCTTCACAAACTAAGCGGAGTTC


         ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
    1142 CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTCAGAATGAAATC
         GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG
                      ^

         1150 NCOI,


         ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
    1202 ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC
         TGGGACTGCGTGGGTCAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG
                                                ^  ^  ^        ^     ^

         1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,


         ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
    1262 GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG
```

FIGURE 18 - Page 3

CAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGCGCATAACGGAC

SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
1322 TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA
AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT
                                                          ^
1369 NAEI,

ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
1382 CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA
GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT
     ^
1385 DRD1,

ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
1442 CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC
GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG

LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
1502 CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA
GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
     ^   ^
1502 PSTI, 1507 TTH3I,

LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
1562 AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACTTCATCAGTGGGATACAATACTTG
TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC
     ^   ^
1565 XHOI, 1586 NDEI,

AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
1622 GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCGA
                     ^   ^
1643 BSTE2, 1677 ALWN1 PVU2,

AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
1682 GCTGTCACCAGCCCACTAACCACTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG
CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCCACC

ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
1742 GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
CACCGACGGGTCGAGCGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA
                                                      ^
1794 ESP1,

GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
1802 GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT
CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA
     ^
1802 KAS1 NARI,

GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
1862 GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC
CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
                 ^   ^
1878 SACI, 1899 BSPH1,

FIGURE 18 - Page 4

```
          ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
    1922  ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC
          TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
                        ^
    1928 TTH3I,


          ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
    1982  GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG
          CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCGTCACGTCACC
                                                    ^
    2004 NAEI, 2017 SMAI XMAI,


          MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
    2042  ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCCACGCACTACGTG
          TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC
                            ^                                   ^
    2067 SMAI XMAI, 2093 DRA3,


          ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
    2102  CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG
          GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC
                        ^                                           ^
    2115 PVU2, 2159 ALWN1,


          LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
    2162  CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC
          GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
                ^                                                   ^
    2164 MST2, 2220 ECON1,


          TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
    2222  TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA
          ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACTCGCTGAAATTCTGGACCGAT


          LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
    2282  AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT
          TTTCGATTCGAGTACGGTGTCGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCATA
                ^              ^          ^
    2285 ESP1, 2300 PVU2, 2310 BAMHI,


          LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
    2342  AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
          TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG


          ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
    2402  ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG
          TGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGTAC
                            ^                   ^         ^^
    2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,


          TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
    2462  TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCCTGCG
          ACCTCACCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC
                            ^                       ^
    2480 ASE1, 2497 APAI,
```

FIGURE 18 - Page 5

```
       ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln
2522   CCGAACTACACGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG
       GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC
                                                  ^

       2553 PSTI,

       ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
2582   GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG
       CACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC
                     ^

       2594 DRA3,

       ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
2642   GTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC
       CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG

       ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
2702   CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTCAGAGTAGGACTCCACGAATACCCG
       GGGACGTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC
                                                               ^

       2757 HGIE2,

       ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
2762   GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC
       CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACTGCAGGTACGAG
                                                            ^

       2809 AAT2,

       ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
2822   ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGATCACCC
       TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCCTAGTGGG
                              ^

       2850 EAG1 XMA3,

       ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys
2882   CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
       GGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG
                  ^                   ^

       2889 BALI, 2903 NHEI,

       ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
2942   ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
       TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC
                                          ^   ^

       2966 ESP1, 2969 SACI,

       GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
3002   GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC
       CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTTCACCACTAAGACCTGAGG

       PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
3062   TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG
       AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC
                                                         ^

       3096 BGL2,

       ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
```

FIGURE 18 - Page 6

3122 CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC
     GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG
                    ^   ^

3143 ALWN1, 3164 EAG1 XMA3,

     ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
3182 CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG
     GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC
                                        ^   ^

3217 HGIE2, 3229 NCOI,

     LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
3242 CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC
     GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

     ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
3302 ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC
     TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG
                                  ^   ^

3332 SACI, 3346 HIND3,

     SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
3362 TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCCTTCTGGC
     AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

     CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
3422 TGCCCCCCCGACTCCGACGCTGAGTCCTATTCCTCCATGCCCCCCCTGGAGGGGGAGCCT
     ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA
                    ^

3437 EAM11051,

     GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
3482 GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG
     CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC
          ^^ ^

3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

     AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
3542 GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC
     CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG
                                              ^   ^

3589 DRA3, 3600 SAC2,

     AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
3602 GCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT
     CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTGGTGTTA
            ^   ^

3611 ALWN1, 3655 PFLM1,

     LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
3662 TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC
     AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAACTG
                        ^

3681 DRA3,

     ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
3722 AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG

FIGURE 18 - Page 7

TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCGTCGCCGC

```
      SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
3782  TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCCACAC
      AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG
                                            ^

3816 HIND3,

      SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
3842  TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC
      AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG
                                        ^                  ^

3875 AAT2, 3890 BGLI,

      ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
3902  GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC
      CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG

      ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
3962  ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG
      TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCCAGCATTC

      ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
4022  CCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG
      GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC

      TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
4082  TACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC
      ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG

      SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
4142  TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG
      AGTGGTCCTGTCGCCCAACTTAAGGAGCACGTTCGCACCTTCAGGTTCTTTTGGGGTTAC
                    ^

4160 ECORI,

      GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
4202  GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
      CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC
                              ^           ^

4229 DRD1, 4236 ALWN1,

      GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
4262  GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCGTGGCCATCAAGTCC
      CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGGTTCGGGCGCACCGGTAGTTCAGG
                                        ^            ^

4301 BGLI, 4308 BALI,

      LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
4322  CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC
      GAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTTCCCCCCTCTTGACGCCG
                            ^

4345 APAI,

      TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys
4382  TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC
      ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG
```

FIGURE 18 - Page 8

```
      TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
4442  TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG
      ATGTAGTTCCGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC
                    ^
      4452 SMAI XMAI,


      CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
4502  TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC
      ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG
            ^    ^
      4508 DRD1, 4511 TTH3I,


      LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
4562  CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCCTGGGGACCCCCCACAA
      GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGGACCCCTGGGGGGTGTT


      ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp
4622  CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC
      GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG
                              ^
      4637 SACI,


      GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
4682  GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT
      CCGCGACCTTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA
                                                        ^
      4731 NRUI,


      AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
4742  GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT
      CGCACCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA


      AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla
4802  GCCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC
      CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG
            ^^
      4806 PFLM1, 4807 DRA3,


      ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu
4862  AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA
      TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT
                                          ^
      4893 BGL2,


      ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
4922  CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC
      GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG
                                        ^
      4954 NCOI,


      SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
4982  AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACTTGGGGTACCG
      TCAATGAGAGGTCCACTTTAGTTATCCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC
                              ^                                 ^
      5015 SPHI, 5035 KPNI,
```

FIGURE 18 - Page 9

```
        ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly
  5042  CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
        GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT
                      ^                                   ^

  5064 APAI, 5091 BALI,


        GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
  5102  GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA
        CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT
                       ^

  5113 NDEI,


        LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
  5162  CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC
        GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG
                       ^^          ^        ^

  5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,


        SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
  5222  AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCCGCTGGATCTGGTTTTGC
        TCGCCCCCTCTGTAAATAGTGTCGCACAGAGTACGGGCCGGGGCGACCTAGACCAAAACG
                     ^

  5240 DRA3,


        LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn
  5282  CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGAATGAGCACGAAT
        GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA
                     ^

  5295 PSTI,


        ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe
  5342  CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCGGCCGCAGGACGTCAAGTTC
        GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGGCGTCCTGCAGTTCAAG
                                                ^^          ^         ^

  5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,


        ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu
  5402  CCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTTGCCGCGCAGGGGCCCTAGATTG
        GGCCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC
                                                                 ^

  5449 APAI,


        GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro
  5462  GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT
        CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA
                 ^            ^                       ^             ^

  5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,


        IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
  5522  ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC
        TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGG
                      ^                      ^        ^   ^

  5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,


        LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
  5582  CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG
        GAGATACCGTTACTCCCGACGCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC
```

FIGURE 18 - Page 10

```
       ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp
5642   CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT
       GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA
                    ^                                            ^

       5650 APAI, 5696 CLAI,


       ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu
5702   ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCCTCTT
       TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA
                    ^                                  ^        ^

       5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,


       GlyGlyAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAspGlyValAsnTyr
5762   GGAGGCGCTGCCAGGGCCCTGGCGCATGGCGTCCGGGTTCTGGAAGACGGCGTGAACTAT
       CCTCCGCGACGGTCCCGGGACCGCGTACCGCAGGCCCAAGACCTTCTGCCGCACTTGATA
                 ^  ^

       5772 BSTXI, 5775 APAI,


       AlaThrGlyAsnLeuProGlyCysSerOC AM
5822   GCAACAGGGAACCTTCCTGGTTGCTCTTAATAGTCGAC
       CGTTGTCCCTTGGAAGGACCAACGAGAATTATCAGCTG
                                       ^

       5854 SALI,
```

FIGURE 19

FIGURE 20 - Page

FIGURE 20 - Pa

pSP72
HindIII-SalI
clone #6

NotI

121bp

PstI

pd.ΔNS3NS5.PJ
core140RT

NotI

PstI

SalI

pd.ΔNS3NS5.PJ
core150RT

NotI

PstI

SalI

Ligate fragments into pd.NS3NS5.PJ
NotI-SalI cloning vector.
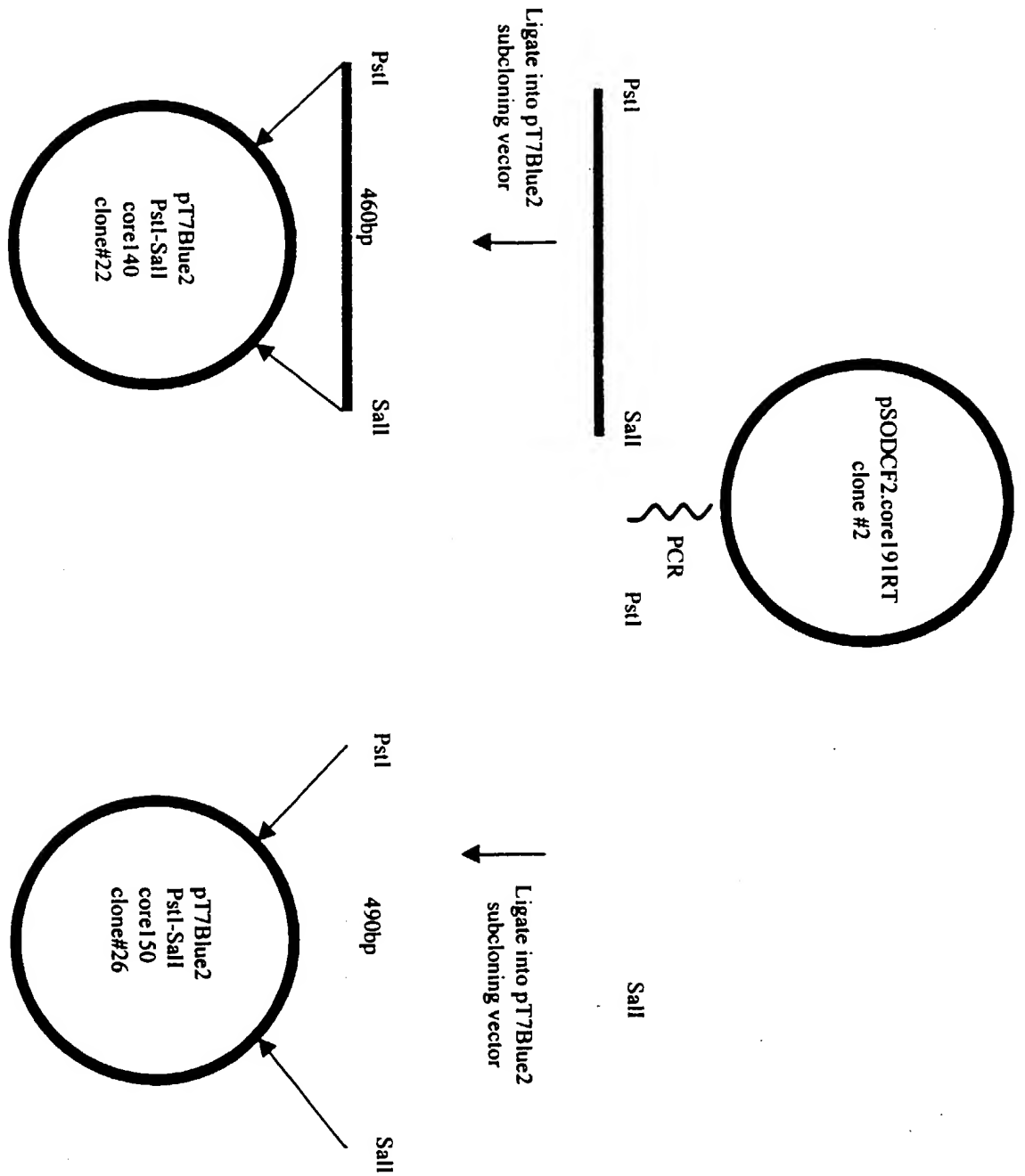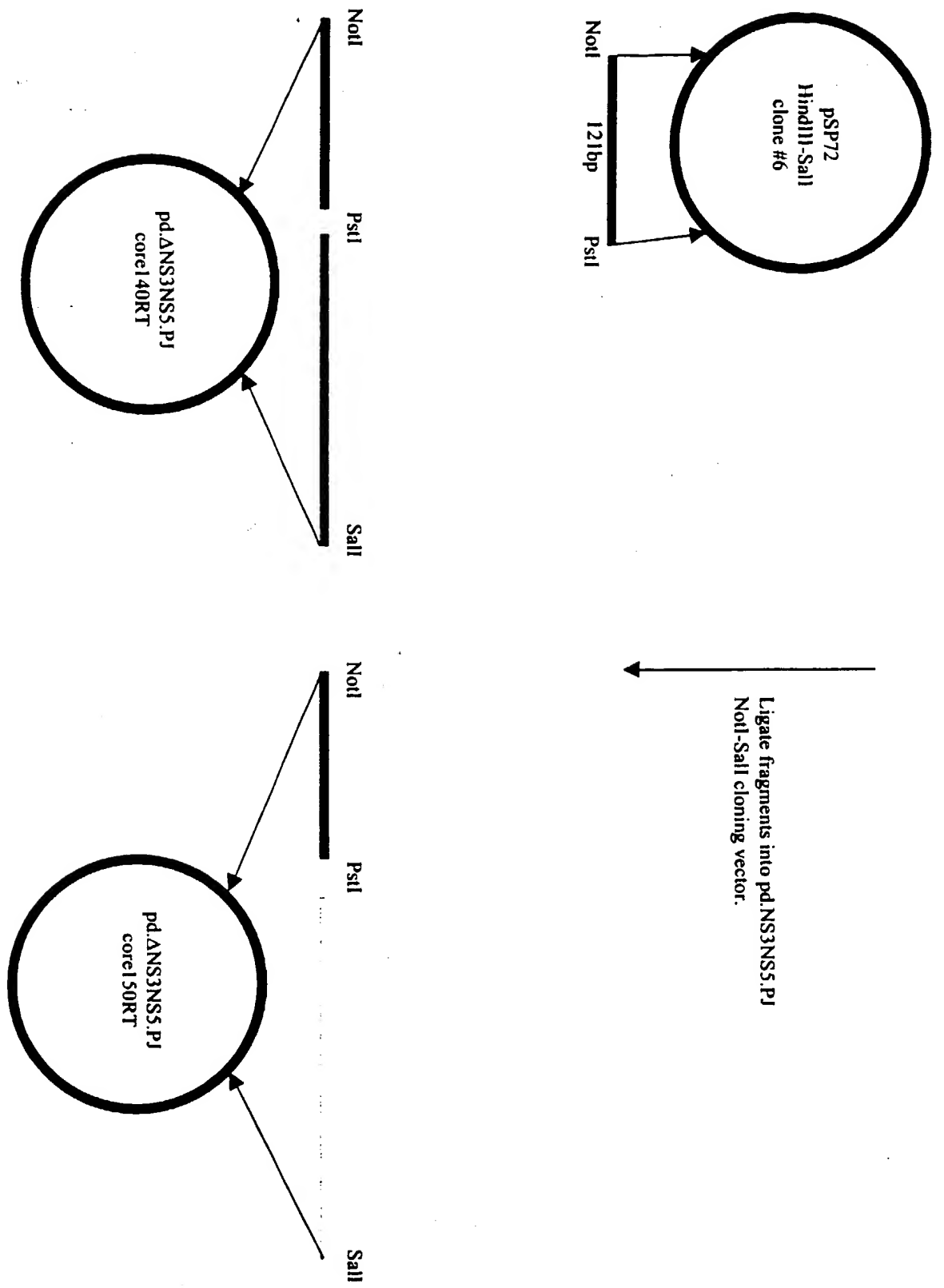
FIGURE 21 - Page 1

```
                   MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn
   2  AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
      TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG
      ^                        ^                          ^

   1 HIND3, 24 NDEI, 52 SCAI,

         ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp
  62  CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT
      GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA
                                                                 ^

  116 CLAI,

         ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr
 122  CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC
      GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG

         TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys
 182  TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT
      ATGCCGTTCAAGGAACGGCTGCCGCCCACGAGCCCCCGCGAATACTGTATTATTAAACA

         AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln
 242  GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA
      CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAACTGGTT

         AlaGluThrAlaGlyAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal
 302  GCAGAGACTGCGGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC
      CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG
                                                                 ^

  303 ALWN1,

         ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe
 362  ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT
      TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA

         TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
 422  TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTCAT
      ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCCTCTGTAGAGTAGAAGACAGTA
```

FIGURE 21 - Page 2

```
          SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
    482   TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG
          AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC


          AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
    542   GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTCGTCGTG
          CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC
                 ^           ^
          550 SAC2, 560 DRD1,


          AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn
    602   GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT
          CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA
                        ^
          615 BSPH1,


          ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle
    662   ACGTGTGTCACCCAGACAGTCGATTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC
          TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAACTGGGATGGAAGTGGTAACTCTGTTAG


          ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
    722   ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG
          TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC


          ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
    782   CCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCCTCCGGCATGTTCGACTCGTCC
          GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG
                                   ^                       ^
          816 BGLI, 833 DRD1,


          ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
    842   GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT
          CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA
                                                       ^
          881 SACI,


          ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
    902   ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT
          TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAA
                                     ^
          931 SMAI XMAI,


          GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
    962   GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG
          CTTAAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC
                                             ^
          985 STUI,


          ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
   1022   ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT
          TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA
                                                         ^
          1069 DRA3,


          ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
   1082   AGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTCGCCTCAAG
```

FIGURE 21 - Page 3

TCCCGAGTTCGGGGAGGGGGTAGCACCCTGGTCTACACCTTCACAAACTAAGCGGAGTTC

```
      ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
1142  CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTCAGAATGAAATC
      GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG
                    ^
1150  NCOI,
```

```
      ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
1202  ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC
      TGGGACTGCGTGGGTCAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG
                        ^     ^   ^         ^      ^
1230  BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,
```

```
      ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
1262  GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG
      CAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGCGCATAACGGAC
```

```
      SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
1322  TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA
      AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT
                                                              ^
1369  NAEI,
```

```
      ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
1382  CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA
      GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT
               ^
1385  DRD1,
```

```
      ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
1442  CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC
      GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG
```

```
      LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
1502  CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA
      GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
      ^       ^
1502  PSTI, 1507 TTH3I,
```

```
      LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
1562  AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACTTCATCAGTGGGATACAATACTTG
      TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC
               ^                          ^
1565  XHOI, 1586 NDEI,
```

```
      AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
1622  GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
      CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCGA
                        ^                                      ^
1643  BSTE2, 1677 ALWN1 PVU2,
```

```
      AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
1682  GCTGTCACCAGCCCACTAACCACTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG
      CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCCACC
```

FIGURE 21 - Page

```
      ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
1742  GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
      CACCGACGGGTCGAGCGGCGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA
                                                               ^

    1794 ESP1,


      GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
1802  GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT
      CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA
      ^

    1802 KAS1 NARI,


      GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
1862  GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC
      CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
                            ^                   ^

    1878 SACI, 1899 BSPH1,


      ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
1922  ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC
      TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
                            ^

    1928 TTH3I,


      ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
1982  GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG
      CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC
                                  ^                 ^

    2004 NAEI, 2017 SMAI XMAI,


      MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
2042  ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCCACGCACTACGTG
      TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC
                              ^                             ^

    2067 SMAI XMAI, 2093 DRA3,


      ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
2102  CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG
      GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC
                        ^                                     ^

    2115 PVU2, 2159 ALWN1,


      LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
2162  CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC
      GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
      ^                                                           ^

    2164 MST2, 2220 ECON1,


      TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
2222  TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA
      ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACTCGCTGAAATTCTGGACCGAT


      LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
2282  AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT
      TTTCGATTCGAGTACGGTGTCGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCATA
      ^                     ^                   ^

    2285 ESP1, 2300 PVU2, 2310 BAMHI,
```

FIGURE 21 - Page 5

```
          LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
    2342  AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCCACTGTGGAGCTGAGATC
          TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG

          ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
    2402  ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG
          TGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGTAC
                                              ^              ^     ^^
    2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,


          TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
    2462  TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCCTGCG
          ACCTCACCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC
                                        ^                    ^
    2480 ASE1, 2497 APAI,


          ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln
    2522  CCGAACTACACGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG
          GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC
                                                    ^
    2553 PSTI,


          ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
    2582  GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG
          CACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC
                     ^
    2594 DRA3,


          ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
    2642  GTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC
          CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG


          ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
    2702  CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTCAGAGTAGGACTCCACGAATACCCG
          GGGACGTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC
                                                                    ^
    2757 HGIE2,


          ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
    2762  GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC
          CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACTGCAGGTACGAG
                                                            ^
    2809 AAT2,


          ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
    2822  ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGATCACCC
          TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCCTAGTGGG
                                              ^
    2850 EAG1 XMA3,


          ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys
    2882  CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
          GGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG
                     ^                    ^
    2889 BALI, 2903 NHEI,
```

FIGURE 21 - Page 6

```
      ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
2942  ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
      TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC
                                       ^   ^

      2966 ESP1, 2969 SACI,

      GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
3002  GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC
      CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTTCACCACTAAGACCTGAGG

      PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
3062  TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG
      AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC
                                          ^

      3096 BGL2,

      ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
3122  CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC
      GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG
                               ^                        ^

      3143 ALWN1, 3164 EAG1 XMA3,

      ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
3182  CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG
      GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC
                                          ^              ^

      3217 HGIE2, 3229 NCOI,

      LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
3242  CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC
      GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG

      ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
3302  ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC
      TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG
                                       ^              ^

      3332 SACI, 3346 HIND3,

      SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
3362  TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCCTTCTGGC
      AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG

      CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
3422  TGCCCCCCCGACTCCGACGCTGAGTCCTATTCCTCCATGCCCCCCCTGGAGGGGGAGCCT
      ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA
                                  ^

      3437 EAM11051,

      GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
3482  GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG
      CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC
                      ^^  ^

      3484 BAMHI, 3485 BSAB1, 3487 BSPE1,

      AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
3542  GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC
      CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG
```

FIGURE 21 - Page

3589 DRA3, 3600 SAC2,

```
       AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
3602   GCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT
       CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTGGTGTTA
          ^                                                       ^
```

3611 ALWN1, 3655 PFLM1,

```
       LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
3662   TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC
       AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAACTG
                          ^
```

3681 DRA3,

```
       ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
3722   AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG
       TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCGTCGCCGC
```

```
       SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
3782   TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCCACAC
       AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG
                                                  ^
```

3816 HIND3,

```
       SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
3842   TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC
       AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG
                                          ^                ^
```

3875 AAT2, 3890 BGLI,

```
       ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
3902   GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC
       CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG
```

```
       ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
3962   ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG
       TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCCAGCATTC
```

```
       ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
4022   CCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG
       GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC
```

```
       TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
4082   TACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC
       ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG
```

```
       SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
4142   TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG
       AGTGGTCCTGTCGCCCAACTTAAGGAGCACGTTCGCACCTTCAGGTTCTTTTGGGGTTAC
                          ^
```

4160 ECORI,

```
       GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
4202   GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
       CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC
                          ^                ^
```

FIGURE 21 - Page 8

4229 DRD1, 4236 ALWN1,

```
       GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
4262   GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCGTGGCCATCAAGTCC
       CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCGGGCGCACCGGTAGTTCAGG
                                                ^         ^
```

4301 BGLI, 4308 BALI,

```
       LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
4322   CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC
       GAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTTCCCCCCTCTTGACGCCG
                              ^
```

4345 APAI,

```
       TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys
4382   TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC
       ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG
```

```
       TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
4442   TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG
       ATGTAGTTCCGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC
                    ^
```

4452 SMAI XMAI,

```
       CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
4502   TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC
       ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG
                ^  ^
```

4508 DRD1, 4511 TTH3I,

```
       LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
4562   CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCCTGGGGACCCCCCACAA
       GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGGACCCCTGGGGGGTGTT
```

```
       ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp
4622   CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC
       GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG
                                        ^
```

4637 SACI,

```
       GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
4682   GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT
       CCGCGACCTTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA
                                                            ^
```

4731 NRUI,

```
       AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
4742   GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT
       CGCACCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA
```

```
       AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla
4802   GCCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC
       CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG
                ^^
```

4806 PFLM1, 4807 DRA3,

```
       ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu
```

FIGURE 21 - Page

4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA
     TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT
                                         ^

     4893 BGL2,

     ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC
     GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG
                                   ^

     4954 NCOI,

     SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACTTGGGGTACCG
     TCAATGAGAGGTCCACTTTAGTTATCCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC
                                      ^   ^

     5015 SPHI, 5035 KPNI,

     ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly
5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
     GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT
                              ^   ^

     5064 APAI, 5091 BALI,

     GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA
     CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT
                        ^

     5113 NDEI,

     LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC
     GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG
                   ^^   ^   ^

     5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

     SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCCGCTGGATCTGGTTTTGC
     TCGCCCCCTCTGTAAATAGTGTCGCACAGAGTACGGGCCGGGGCGACCTAGACCAAAACG
                     ^

     5240 DRA3,

     LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGAATGAGCACGAAT
     GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA
                   ^

     5295 PSTI,

     ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe
5342 CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCGGCCGCAGGACGTCAAGTTC
     GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGGCGTCCTGCAGTTCAAG
                                             ^^   ^   ^

     5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

     ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu
5402 CCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTTGCCGCGCAGGGGCCCTAGATTG
     GGCCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC
                                              ^

FIGURE 21 - Page 10

5449 APAI,

```
       GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro
5462   GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT
       CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA
              ^               ^                  ^          ^
```

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

```
       IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
5522   ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC
       TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGG
                                    ^         ^     ^    ^
```

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

```
       LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
5582   CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG
       GAGATACCGTTACTCCCGACGCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC
```

```
       ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp
5642   CCTAGCTGGGGCCCCACAGACCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT
       GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA
              ^                                            ^
```

5650 APAI, 5696 CLAI,

```
       ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValOC AM
5702   ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCTAATAGTCGAC
       TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGATTATCAGCTG
              ^                                            ^
```

5724 HGIE2, 5755 SALI,

FIGURE 22 - Page 1

```
                    MetAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn
   2   AGCTTACAAAACAAAATGGCTGCATATGCAGCTCAGGGCTATAAGGTGCTAGTACTCAAC
       TCGAATGTTTTGTTTTACCGACGTATACGTCGAGTCCCGATATTCCACGATCATGAGTTG
       ^                   ^                                    ^
   1 HIND3, 24 NDEI, 52 SCAI,


       ProSerValAlaAlaThrLeuGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleAsp
  62   CCCTCTGTTGCTGCAACACTGGGCTTTGGTGCTTACATGTCCAAGGCTCATGGGATCGAT
       GGGAGACAACGACGTTGTGACCCGAAACCACGAATGTACAGGTTCCGAGTACCCTAGCTA
                                                                 ^
       116 CLAI,


       ProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerThr
  122  CCTAACATCAGGACCGGGGTGAGAACAATTACCACTGGCAGCCCCATCACGTACTCCACC
       GGATTGTAGTCCTGGCCCCACTCTTGTTAATGGTGACCGTCGGGGTAGTGCATGAGGTGG


       TyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyAlaTyrAspIleIleIleCys
  182  TACGGCAAGTTCCTTGCCGACGGCGGGTGCTCGGGGGGCGCTTATGACATAATAATTTGT
       ATGCCGTTCAAGGAACGGCTGCCGCCCACGAGCCCCCCGCGAATACTGTATTATTAAACA


       AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln
  242  GACGAGTGCCACTCCACGGATGCCACATCCATCTTGGGCATTGGCACTGTCCTTGACCAA
       CTGCTCACGGTGAGGTGCCTACGGTGTAGGTAGAACCCGTAACCGTGACAGGAACTGGTT


       AlaGluThrAlaGlyAlaAlaArgLeuValValLeuAlaThrAlaThrProProGlySerVal
  302  GCAGAGACTGCGGGGGCGAGACTGGTTGTGCTCGCCACCGCCACCCCTCCGGGCTCCGTC
       CGTCTCTGACGCCCCCGCTCTGACCAACACGAGCGGTGGCGGTGGGGAGGCCCGAGGCAG
       ^
       303 ALWN1,


       ThrValProHisProAsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPhe
  362  ACTGTGCCCCATCCCAACATCGAGGAGGTTGCTCTGTCCACCACCGGAGAGATCCCTTTT
       TGACACGGGGTAGGGTTGTAGCTCCTCCAACGAGACAGGTGGTGGCCTCTCTAGGGAAAA


       TyrGlyLysAlaIleProLeuGluValIleLysGlyGlyArgHisLeuIlePheCysHis
  422  TACGGCAAGGCTATCCCCCTCGAAGTAATCAAGGGGGGGAGACATCTCATCTTCTGTCAT
       ATGCCGTTCCGATAGGGGGAGCTTCATTAGTTCCCCCCCTCTGTAGAGTAGAAGACAGTA
```

FIGURE 22 - Page 2

```
      SerLysLysLysCysAspGluLeuAlaAlaLysLeuValAlaLeuGlyIleAsnAlaVal
482   TCAAAGAAGAAGTGCGACGAACTCGCCGCAAAGCTGGTCGCATTGGGCATCAATGCCGTG
      AGTTTCTTCTTCACGCTGCTTGAGCGGCGTTTCGACCAGCGTAACCCGTAGTTACGGCAC


      AlaTyrTyrArgGlyLeuAspValSerValIleProThrSerGlyAspValValValVal
542   GCCTACTACCGCGGTCTTGACGTGTCCGTCATCCCGACCAGCGGCGATGTTGTCGTCGTG
      CGGATGATGGCGCCAGAACTGCACAGGCAGTAGGGCTGGTCGCCGCTACAACAGCAGCAC
                ^                ^
      550 SAC2, 560 DRD1,


      AlaThrAspAlaLeuMetThrGlyTyrThrGlyAspPheAspSerValIleAspCysAsn
602   GCAACCGATGCCCTCATGACCGGCTATACCGGCGACTTCGACTCGGTGATAGACTGCAAT
      CGTTGGCTACGGGAGTACTGGCCGATATGGCCGCTGAAGCTGAGCCACTATCTGACGTTA
                       ^
      615 BSPH1,


      ThrCysValThrGlnThrValAspPheSerLeuAspProThrPheThrIleGluThrIle
662   ACGTGTGTCACCCAGACAGTCGATTTCAGCCTTGACCCTACCTTCACCATTGAGACAATC
      TGCACACAGTGGGTCTGTCAGCTAAAGTCGGAACTGGGATGGAAGTGGTAACTCTGTTAG


      ThrLeuProGlnAspAlaValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLys
722   ACGCTCCCCCAAGATGCTGTCTCCCGCACTCAACGTCGGGGCAGGACTGGCAGGGGGAAG
      TGCGAGGGGGTTCTACGACAGAGGGCGTGAGTTGCAGCCCCGTCCTGACCGTCCCCCTTC


      ProGlyIleTyrArgPheValAlaProGlyGluArgProSerGlyMetPheAspSerSer
782   CCAGGCATCTACAGATTTGTGGCACCGGGGGAGCGCCCCTCCGGCATGTTCGACTCGTCC
      GGTCCGTAGATGTCTAAACACCGTGGCCCCCTCGCGGGGAGGCCGTACAAGCTGAGCAGG
                                      ^                ^
      816 BGLI, 833 DRD1,


      ValLeuCysGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThr
842   GTCCTCTGTGAGTGCTATGACGCAGGCTGTGCTTGGTATGAGCTCACGCCCGCCGAGACT
      CAGGAGACACTCACGATACTGCGTCCGACACGAACCATACTCGAGTGCGGGCGGCTCTGA
                                    ^
      881 SACI,


      ThrValArgLeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeu
902   ACAGTTAGGCTACGAGCGTACATGAACACCCCGGGGCTTCCCGTGTGCCAGGACCATCTT
      TGTCAATCCGATGCTCGCATGTACTTGTGGGGCCCCGAAGGGCACACGGTCCTGGTAGAA
                                ^
      931 SMAI XMAI,


      GluPheTrpGluGlyValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGln
962   GAATTTTGGGAGGGCGTCTTTACAGGCCTCACTCATATAGATGCCCACTTTCTATCCCAG
      CTTAAAACCCTCCCGCAGAAATGTCCGGAGTGAGTATATCTACGGGTGAAAGATAGGGTC
                                       ^
      985 STUI,


      ThrLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAla
1022  ACAAAGCAGAGTGGGGAGAACCTTCCTTACCTGGTAGCGTACCAAGCCACCGTGTGCGCT
      TGTTTCGTCTCACCCCTCTTGGAAGGAATGGACCATCGCATGGTTCGGTGGCACACGCGA
                                                        ^
      1069 DRA3,


      ArgAlaGlnAlaProProProSerTrpAspGlnMetTrpLysCysLeuIleArgLeuLys
1082  AGGGCTCAAGCCCCTCCCCCATCGTGGGACCAGATGTGGAAGTGTTTGATTCGCCTCAAG
```

FIGURE 22 - Page

TCCCGAGTTCGGGGAGGGGGTAGCACCCTGGTCTACACCTTCACAAACTAAGCGGAGTTC

```
        ProThrLeuHisGlyProThrProLeuLeuTyrArgLeuGlyAlaValGlnAsnGluIle
1142    CCCACCCTCCATGGGCCAACACCCCTGCTATACAGACTGGGCGCTGTTCAGAATGAAATC
        GGGTGGGAGGTACCCGGTTGTGGGGACGATATGTCTGACCCGCGACAAGTCTTACTTTAG
                    ^
1150 NCOI,
```

```
        ThrLeuThrHisProValThrLysTyrIleMetThrCysMetSerAlaAspLeuGluVal
1202    ACCCTGACGCACCCAGTCACCAAATACATCATGACATGCATGTCGGCCGACCTGGAGGTC
        TGGGACTGCGTGGGTCAGTGGTTTATGTAGTACTGTACGTACAGCCGGCTGGACCTCCAG
                  ^     ^   ^           ^       ^
1230 BSPH1, 1234 DRD1, 1237 AVA3, 1245 EAG1 XMA3, 1250 DRD1,
```

```
        ValThrSerThrTrpValLeuValGlyGlyValLeuAlaAlaLeuAlaAlaTyrCysLeu
1262    GTCACGAGCACCTGGGTGCTCGTTGGCGGCGTCCTGGCTGCTTTGGCCGCGTATTGCCTG
        CAGTGCTCGTGGACCCACGAGCAACCGCCGCAGGACCGACGAAACCGGCGCATAACGGAC
```

```
        SerThrGlyCysValValIleValGlyArgValValLeuSerGlyLysProAlaIleIle
1322    TCAACAGGCTGCGTGGTCATAGTGGGCAGGGTCGTCTTGTCCGGGAAGCCGGCAATCATA
        AGTTGTCCGACGCACCAGTATCACCCGTCCCAGCAGAACAGGCCCTTCGGCCGTTAGTAT
                                                            ^
1369 NAEI,
```

```
        ProAspArgGluValLeuTyrArgGluPheAspGluMetGluGluCysSerGlnHisLeu
1382    CCTGACAGGGAAGTCCTCTACCGAGAGTTCGATGAGATGGAAGAGTGCTCTCAGCACTTA
        GGACTGTCCCTTCAGGAGATGGCTCTCAAGCTACTCTACCTTCTCACGAGAGTCGTGAAT
              ^
1385 DRD1,
```

```
        ProTyrIleGluGlnGlyMetMetLeuAlaGluGlnPheLysGlnLysAlaLeuGlyLeu
1442    CCGTACATCGAGCAAGGGATGATGCTCGCCGAGCAGTTCAAGCAGAAGGCCCTCGGCCTC
        GGCATGTAGCTCGTTCCCTACTACGAGCGGCTCGTCAAGTTCGTCTTCCGGGAGCCGGAG
```

```
        LeuGlnThrAlaSerArgGlnAlaGluValIleAlaProAlaValGlnThrAsnTrpGln
1502    CTGCAGACCGCGTCCCGTCAGGCAGAGGTTATCGCCCCTGCTGTCCAGACCAACTGGCAA
        GACGTCTGGCGCAGGGCAGTCCGTCTCCAATAGCGGGGACGACAGGTCTGGTTGACCGTT
        ^    ^
1502 PSTI, 1507 TTH3I,
```

```
        LysLeuGluThrPheTrpAlaLysHisMetTrpAsnPheIleSerGlyIleGlnTyrLeu
1562    AAACTCGAGACCTTCTGGGCGAAGCATATGTGGAACTTCATCAGTGGGATACAATACTTG
        TTTGAGCTCTGGAAGACCCGCTTCGTATACACCTTGAAGTAGTCACCCTATGTTATGAAC
            ^                       ^
1565 XHOI, 1586 NDEI,
```

```
        AlaGlyLeuSerThrLeuProGlyAsnProAlaIleAlaSerLeuMetAlaPheThrAla
1622    GCGGGCTTGTCAACGCTGCCTGGTAACCCCGCCATTGCTTCATTGATGGCTTTTACAGCT
        CGCCCGAACAGTTGCGACGGACCATTGGGGCGGTAACGAAGTAACTACCGAAAATGTCGA
                  ^                                             ^
1643 BSTE2, 1677 ALWN1 PVU2,
```

```
        AlaValThrSerProLeuThrThrSerGlnThrLeuLeuPheAsnIleLeuGlyGlyTrp
1682    GCTGTCACCAGCCCACTAACCACTAGCCAAACCCTCCTCTTCAACATATTGGGGGGGTGG
        CGACAGTGGTCGGGTGATTGGTGATCGGTTTGGGAGGAGAAGTTGTATAACCCCCCCACC
```

FIGURE 22 - Page 4

```
        ValAlaAlaGlnLeuAlaAlaProGlyAlaAlaThrAlaPheValGlyAlaGlyLeuAla
1742    GTGGCTGCCCAGCTCGCCGCCCCCGGTGCCGCTACTGCCTTTGTGGGCGCTGGCTTAGCT
        CACCGACGGGTCGAGCGGCGGGGGGCCACGGCGATGACGGAAACACCCGCGACCGAATCGA
                                                                   ^

    1794 ESP1,


        GlyAlaAlaIleGlySerValGlyLeuGlyLysValLeuIleAspIleLeuAlaGlyTyr
1802    GGCGCCGCCATCGGCAGTGTTGGACTGGGGAAGGTCCTCATAGACATCCTTGCAGGGTAT
        CCGCGGCGGTAGCCGTCACAACCTGACCCCTTCCAGGAGTATCTGTAGGAACGTCCCATA
        ^

    1802 KAS1 NARI,


        GlyAlaGlyValAlaGlyAlaLeuValAlaPheLysIleMetSerGlyGluValProSer
1862    GGCGCGGGCGTGGCGGGAGCTCTTGTGGCATTCAAGATCATGAGCGGTGAGGTCCCCTCC
        CCGCGCCCGCACCGCCCTCGAGAACACCGTAAGTTCTAGTACTCGCCACTCCAGGGGAGG
                                ^                   ^

    1878 SACI, 1899 BSPH1,


        ThrGluAspLeuValAsnLeuLeuProAlaIleLeuSerProGlyAlaLeuValValGly
1922    ACGGAGGACCTGGTCAATCTACTGCCCGCCATCCTCTCGCCCGGAGCCCTCGTAGTCGGC
        TGCCTCCTGGACCAGTTAGATGACGGGCGGTAGGAGAGCGGGCCTCGGGAGCATCAGCCG
                            ^

    1928 TTH3I,


        ValValCysAlaAlaIleLeuArgArgHisValGlyProGlyGluGlyAlaValGlnTrp
1982    GTGGTCTGTGCAGCAATACTGCGCCGGCACGTTGGCCCGGGCGAGGGGGCAGTGCAGTGG
        CACCAGACACGTCGTTATGACGCGGCCGTGCAACCGGGCCCGCTCCCCCGTCACGTCACC
                            ^               ^

    2004 NAEI, 2017 SMAI XMAI,


        MetAsnArgLeuIleAlaPheAlaSerArgGlyAsnHisValSerProThrHisTyrVal
2042    ATGAACCGGCTGATAGCCTTCGCCTCCCGGGGGAACCATGTTTCCCCCACGCACTACGTG
        TACTTGGCCGACTATCGGAAGCGGAGGGCCCCCTTGGTACAAAGGGGGTGCGTGATGCAC
                            ^                           ^

    2067 SMAI XMAI, 2093 DRA3,


        ProGluSerAspAlaAlaAlaArgValThrAlaIleLeuSerSerLeuThrValThrGln
2102    CCGGAGAGCGATGCAGCTGCCCGCGTCACTGCCATACTCAGCAGCCTCACTGTAACCCAG
        GGCCTCTCGCTACGTCGACGGGCGCAGTGACGGTATGAGTCGTCGGAGTGACATTGGGTC
                    ^                                       ^

    2115 PVU2, 2159 ALWN1,


        LeuLeuArgArgLeuHisGlnTrpIleSerSerGluCysThrThrProCysSerGlySer
2162    CTCCTGAGGCGACTGCACCAGTGGATAAGCTCGGAGTGTACCACTCCATGCTCCGGTTCC
        GAGGACTCCGCTGACGTGGTCACCTATTCGAGCCTCACATGGTGAGGTACGAGGCCAAGG
        ^                                                          ^

    2164 MST2, 2220 ECON1,


        TrpLeuArgAspIleTrpAspTrpIleCysGluValLeuSerAspPheLysThrTrpLeu
2222    TGGCTAAGGGACATCTGGGACTGGATATGCGAGGTGTTGAGCGACTTTAAGACCTGGCTA
        ACCGATTCCCTGTAGACCCTGACCTATACGCTCCACAACTCGCTGAAATTCTGGACCGAT


        LysAlaLysLeuMetProGlnLeuProGlyIleProPheValSerCysGlnArgGlyTyr
2282    AAAGCTAAGCTCATGCCACAGCTGCCTGGGATCCCCTTTGTGTCCTGCCAGCGCGGGTAT
        TTTCGATTCGAGTACGGTGTCGACGGACCCTAGGGGAAACACAGGACGGTCGCGCCCATA
        ^                   ^               ^

    2285 ESP1, 2300 PVU2, 2310 BAMHI,
```

FIGURE 22 - Page 5

```
        LysGlyValTrpArgGlyAspGlyIleMetHisThrArgCysHisCysGlyAlaGluIle
2342    AAGGGGGTCTGGCGAGGGGACGGCATCATGCACACTCGCTGCCACTGTGGAGCTGAGATC
        TTCCCCCAGACCGCTCCCCTGCCGTAGTACGTGTGAGCGACGGTGACACCTCGACTCTAG
```

```
        ThrGlyHisValLysAsnGlyThrMetArgIleValGlyProArgThrCysArgAsnMet
2402    ACTGGACATGTCAAAAACGGGACGATGAGGATCGTCGGTCCTAGGACCTGCAGGAACATG
        TGACCTGTACAGTTTTTGCCCTGCTACTCCTAGCAGCCAGGATCCTGGACGTCCTTGTAC
                                    ^                   ^   ^^
```

        2425 BSAB1, 2441 AVR2, 2448 SSE83871, 2449 PSTI,

```
        TrpSerGlyThrPheProIleAsnAlaTyrThrThrGlyProCysThrProLeuProAla
2462    TGGAGTGGGACCTTCCCCATTAATGCCTACACCACGGGCCCCTGTACCCCCCTTCCTGCG
        ACCTCACCCTGGAAGGGGTAATTACGGATGTGGTGCCCGGGGACATGGGGGGAAGGACGC
                                ^                   ^
```

        2480 ASE1, 2497 APAI,

```
        ProAsnTyrThrPheAlaLeuTrpArgValSerAlaGluGluTyrValGluIleArgGln
2522    CCGAACTACACGTTCGCGCTATGGAGGGTGTCTGCAGAGGAATACGTGGAGATAAGGCAG
        GGCTTGATGTGCAAGCGCGATACCTCCCACAGACGTCTCCTTATGCACCTCTATTCCGTC
                                                    ^
```

        2553 PSTI,

```
        ValGlyAspPheHisTyrValThrGlyMetThrThrAspAsnLeuLysCysProCysGln
2582    GTGGGGGACTTCCACTACGTGACGGGTATGACTACTGACAATCTTAAATGCCCGTGCCAG
        CACCCCCTGAAGGTGATGCACTGCCCATACTGATGACTGTTAGAATTTACGGGCACGGTC
                            ^
```

        2594 DRA3,

```
        ValProSerProGluPhePheThrGluLeuAspGlyValArgLeuHisArgPheAlaPro
2642    GTCCCATCGCCCGAATTTTTCACAGAATTGGACGGGGTGCGCCTACATAGGTTTGCGCCC
        CAGGGTAGCGGGCTTAAAAAGTGTCTTAACCTGCCCCACGCGGATGTATCCAAACGCGGG
```

```
        ProCysLysProLeuLeuArgGluGluValSerPheArgValGlyLeuHisGluTyrPro
2702    CCCTGCAAGCCCTTGCTGCGGGAGGAGGTATCATTCAGAGTAGGACTCCACGAATACCCG
        GGGACGTTCGGGAACGACGCCCTCCTCCATAGTAAGTCTCATCCTGAGGTGCTTATGGGC
                                                                ^
```

        2757 HGIE2,

```
        ValGlySerGlnLeuProCysGluProGluProAspValAlaValLeuThrSerMetLeu
2762    GTAGGGTCGCAATTACCTTGCGAGCCCGAACCGGACGTGGCCGTGTTGACGTCCATGCTC
        CATCCCAGCGTTAATGGAACGCTCGGGCTTGGCCTGCACCGGCACAACTGCAGGTACGAG
                                                            ^
```

        2809 AAT2,

```
        ThrAspProSerHisIleThrAlaGluAlaAlaGlyArgArgLeuAlaArgGlySerPro
2822    ACTGATCCCTCCCATATAACAGCAGAGGCGGCCGGGCGAAGGTTGGCGAGGGGATCACCC
        TGACTAGGGAGGGTATATTGTCGTCTCCGCCGGCCCGCTTCCAACCGCTCCCCTAGTGGG
                                                    ^
```

        2850 EAG1 XMA3,

```
        ProSerValAlaSerSerSerAlaSerGlnLeuSerAlaProSerLeuLysAlaThrCys
2882    CCCTCTGTGGCCAGCTCCTCGGCTAGCCAGCTATCCGCTCCATCTCTCAAGGCAACTTGC
        GGGAGACACCGGTCGAGGAGCCGATCGGTCGATAGGCGAGGTAGAGAGTTCCGTTGAACG
                        ^               ^
```

        2889 BALI, 2903 NHEI,

FIGURE 22 - Page 6

```
        ThrAlaAsnHisAspSerProAspAlaGluLeuIleGluAlaAsnLeuLeuTrpArgGln
2942    ACCGCTAACCATGACTCCCCTGATGCTGAGCTCATAGAGGCCAACCTCCTATGGAGGCAG
        TGGCGATTGGTACTGAGGGGACTACGACTCGAGTATCTCCGGTTGGAGGATACCTCCGTC
                                            ^   ^

        2966 ESP1, 2969 SACI,


        GluMetGlyGlyAsnIleThrArgValGluSerGluAsnLysValValIleLeuAspSer
3002    GAGATGGGCGGCAACATCACCAGGGTTGAGTCAGAAAACAAAGTGGTGATTCTGGACTCC
        CTCTACCCGCCGTTGTAGTGGTCCCAACTCAGTCTTTTGTTTCACCACTAAGACCTGAGG


        PheAspProLeuValAlaGluGluAspGluArgGluIleSerValProAlaGluIleLeu
3062    TTCGATCCGCTTGTGGCGGAGGAGGACGAGCGGGAGATCTCCGTACCCGCAGAAATCCTG
        AAGCTAGGCGAACACCGCCTCCTCCTGCTCGCCCTCTAGAGGCATGGGCGTCTTTAGGAC
                                                ^

        3096 BGL2,


        ArgLysSerArgArgPheAlaGlnAlaLeuProValTrpAlaArgProAspTyrAsnPro
3122    CGGAAGTCTCGGAGATTCGCCCAGGCCCTGCCCGTTTGGGCGCGGCCGGACTATAACCCC
        GCCTTCAGAGCCTCTAAGCGGGTCCGGGACGGGCAAACCCGCGCCGGCCTGATATTGGGG
                            ^                           ^

        3143 ALWN1, 3164 EAG1 XMA3,


        ProLeuValGluThrTrpLysLysProAspTyrGluProProValValHisGlyCysPro
3182    CCGCTAGTGGAGACGTGGAAAAAGCCCGACTACGAACCACCTGTGGTCCATGGCTGCCCG
        GGCGATCACCTCTGCACCTTTTTCGGGCTGATGCTTGGTGGACACCAGGTACCGACGGGC
                                                    ^                   ^

        3217 HGIE2, 3229 NCOI,


        LeuProProProLysSerProProValProProProArgLysLysArgThrValValLeu
3242    CTTCCACCTCCAAAGTCCCCTCCTGTGCCTCCGCCTCGGAAGAAGCGGACGGTGGTCCTC
        GAAGGTGGAGGTTTCAGGGGAGGACACGGAGGCGGAGCCTTCTTCGCCTGCCACCAGGAG


        ThrGluSerThrLeuSerThrAlaLeuAlaGluLeuAlaThrArgSerPheGlySerSer
3302    ACTGAATCAACCCTATCTACTGCCTTGGCCGAGCTCGCCACCAGAAGCTTTGGCAGCTCC
        TGACTTAGTTGGGATAGATGACGGAACCGGCTCGAGCGGTGGTCTTCGAAACCGTCGAGG
                                            ^                   ^

        3332 SACI, 3346 HIND3,


        SerThrSerGlyIleThrGlyAspAsnThrThrThrSerSerGluProAlaProSerGly
3362    TCAACTTCCGGCATTACGGGCGACAATACGACAACATCCTCTGAGCCCGCCCCTTCTGGC
        AGTTGAAGGCCGTAATGCCCGCTGTTATGCTGTTGTAGGAGACTCGGGCGGGGAAGACCG


        CysProProAspSerAspAlaGluSerTyrSerSerMetProProLeuGluGlyGluPro
3422    TGCCCCCCCGACTCCGACGCTGAGTCCTATTCCTCCATGCCCCCCCTGGAGGGGGAGCCT
        ACGGGGGGGCTGAGGCTGCGACTCAGGATAAGGAGGTACGGGGGGGACCTCCCCCTCGGA
                            ^

        3437 EAM11051,


        GlyAspProAspLeuSerAspGlySerTrpSerThrValSerSerGluAlaAsnAlaGlu
3482    GGGGATCCGGATCTTAGCGACGGGTCATGGTCAACGGTCAGTAGTGAGGCCAACGCGGAG
        CCCCTAGGCCTAGAATCGCTGCCCAGTACCAGTTGCCAGTCATCACTCCGGTTGCGCCTC
        ^^  ^

        3484 BAMHI, 3485 BSAB1, 3487 BSPE1,


        AspValValCysCysSerMetSerTyrSerTrpThrGlyAlaLeuValThrProCysAla
3542    GATGTCGTGTGCTGCTCAATGTCTTACTCTTGGACAGGCGCACTCGTCACCCCGTGCGCC
        CTACAGCACACGACGAGTTACAGAATGAGAACCTGTCCGCGTGAGCAGTGGGGCACGCGG
```

FIGURE 22 - Page 7

3589 DRA3, 3600 SAC2,

```
       AlaGluGluGlnLysLeuProIleAsnAlaLeuSerAsnSerLeuLeuArgHisHisAsn
3602   GCGGAAGAACAGAAACTGCCCATCAATGCACTAAGCAACTCGTTGCTACGTCACCACAAT
       CGCCTTCTTGTCTTTGACGGGTAGTTACGTGATTCGTTGAGCAACGATGCAGTGGTGTTA
```

3611 ALWN1, 3655 PFLM1,

```
       LeuValTyrSerThrThrSerArgSerAlaCysGlnArgGlnLysLysValThrPheAsp
3662   TTGGTGTATTCCACCACCTCACGCAGTGCTTGCCAAAGGCAGAAGAAAGTCACATTTGAC
       AACCACATAAGGTGGTGGAGTGCGTCACGAACGGTTTCCGTCTTCTTTCAGTGTAAACTG
```

3681 DRA3,

```
       ArgLeuGlnValLeuAspSerHisTyrGlnAspValLeuLysGluValLysAlaAlaAla
3722   AGACTGCAAGTTCTGGACAGCCATTACCAGGACGTACTCAAGGAGGTTAAAGCAGCGGCG
       TCTGACGTTCAAGACCTGTCGGTAATGGTCCTGCATGAGTTCCTCCAATTTCGTCGCCGC
```

```
       SerLysValLysAlaAsnLeuLeuSerValGluGluAlaCysSerLeuThrProProHis
3782   TCAAAAGTGAAGGCTAACTTGCTATCCGTAGAGGAAGCTTGCAGCCTGACGCCCCCACAC
       AGTTTTCACTTCCGATTGAACGATAGGCATCTCCTTCGAACGTCGGACTGCGGGGGTGTG
```

3816 HIND3,

```
       SerAlaLysSerLysPheGlyTyrGlyAlaLysAspValArgCysHisAlaArgLysAla
3842   TCAGCCAAATCCAAGTTTGGTTATGGGGCAAAAGACGTCCGTTGCCATGCCAGAAAGGCC
       AGTCGGTTTAGGTTCAAACCAATACCCCGTTTTCTGCAGGCAACGGTACGGTCTTTCCGG
```

3875 AAT2, 3890 BGLI,

```
       ValThrHisIleAsnSerValTrpLysAspLeuLeuGluAspAsnValThrProIleAsp
3902   GTAACCCACATCAACTCCGTGTGGAAAGACCTTCTGGAAGACAATGTAACACCAATAGAC
       CATTGGGTGTAGTTGAGGCACACCTTTCTGGAAGACCTTCTGTTACATTGTGGTTATCTG
```

```
       ThrThrIleMetAlaLysAsnGluValPheCysValGlnProGluLysGlyGlyArgLys
3962   ACTACCATCATGGCTAAGAACGAGGTTTTCTGCGTTCAGCCTGAGAAGGGGGGTCGTAAG
       TGATGGTAGTACCGATTCTTGCTCCAAAAGACGCAAGTCGGACTCTTCCCCCCAGCATTC
```

```
       ProAlaArgLeuIleValPheProAspLeuGlyValArgValCysGluLysMetAlaLeu
4022   CCAGCTCGTCTCATCGTGTTCCCCGATCTGGGCGTGCGCGTGTGCGAAAAGATGGCTTTG
       GGTCGAGCAGAGTAGCACAAGGGGCTAGACCCGCACGCGCACACGCTTTTCTACCGAAAC
```

```
       TyrAspValValThrLysLeuProLeuAlaValMetGlySerSerTyrGlyPheGlnTyr
4082   TACGACGTGGTTACAAAGCTCCCCTTGGCCGTGATGGGAAGCTCCTACGGATTCCAATAC
       ATGCTGCACCAATGTTTCGAGGGGAACCGGCACTACCCTTCGAGGATGCCTAAGGTTATG
```

```
       SerProGlyGlnArgValGluPheLeuValGlnAlaTrpLysSerLysLysThrProMet
4142   TCACCAGGACAGCGGGTTGAATTCCTCGTGCAAGCGTGGAAGTCCAAGAAAACCCCAATG
       AGTGGTCCTGTCGCCCAACTTAAGGAGCACGTTCGCACCTTCAGGTTCTTTTGGGGTTAC
```

4160 ECORI,

```
       GlyPheSerTyrAspThrArgCysPheAspSerThrValThrGluSerAspIleArgThr
4202   GGGTTCTCGTATGATACCCGCTGCTTTGACTCCACAGTCACTGAGAGCGACATCCGTACG
       CCCAAGAGCATACTATGGGCGACGAAACTGAGGTGTCAGTGACTCTCGCTGTAGGCATGC
```

FIGURE 22 - Page

4229 DRD1, 4236 ALWN1,

```
      GluGluAlaIleTyrGlnCysCysAspLeuAspProGlnAlaArgValAlaIleLysSer
4262  GAGGAGGCAATCTACCAATGTTGTGACCTCGACCCCCAAGCCCGCGTGGCCATCAAGTCC
      CTCCTCCGTTAGATGGTTACAACACTGGAGCTGGGGGTTCGGGCGCACCGGTAGTTCAGG
                                            ^                ^
```

4301 BGLI, 4308 BALI,

```
      LeuThrGluArgLeuTyrValGlyGlyProLeuThrAsnSerArgGlyGluAsnCysGly
4322  CTCACCGAGAGGCTTTATGTTGGGGGCCCTCTTACCAATTCAAGGGGGGAGAACTGCGGC
      GAGTGGCTCTCCGAAATACAACCCCCGGGAGAATGGTTAAGTTCCCCCCTCTTGACGCCG
                                   ^
```

4345 APAI,

```
      TyrArgArgCysArgAlaSerGlyValLeuThrThrSerCysGlyAsnThrLeuThrCys
4382  TATCGCAGGTGCCGCGCGAGCGGCGTACTGACAACTAGCTGTGGTAACACCCTCACTTGC
      ATAGCGTCCACGGCGCGCTCGCCGCATGACTGTTGATCGACACCATTGTGGGAGTGAACG
```

```
      TyrIleLysAlaArgAlaAlaCysArgAlaAlaGlyLeuGlnAspCysThrMetLeuVal
4442  TACATCAAGGCCCGGGCAGCCTGTCGAGCCGCAGGGCTCCAGGACTGCACCATGCTCGTG
      ATGTAGTTCCGGGCCCGTCGGACAGCTCGGCGTCCCGAGGTCCTGACGTGGTACGAGCAC
                      ^
```

4452 SMAI XMAI,

```
      CysGlyAspAspLeuValValIleCysGluSerAlaGlyValGlnGluAspAlaAlaSer
4502  TGTGGCGACGACTTAGTCGTTATCTGTGAAAGCGCGGGGGTCCAGGAGGACGCGGCGAGC
      ACACCGCTGCTGAATCAGCAATAGACACTTTCGCGCCCCCAGGTCCTCCTGCGCCGCTCG
                   ^  ^
```

4508 DRD1, 4511 TTH3I,

```
      LeuArgAlaPheThrGluAlaMetThrArgTyrSerAlaProProGlyAspProProGln
4562  CTGAGAGCCTTCACGGAGGCTATGACCAGGTACTCCGCCCCCCCTGGGGACCCCCCACAA
      GACTCTCGGAAGTGCCTCCGATACTGGTCCATGAGGCGGGGGGGACCCCTGGGGGGTGTT
```

```
      ProGluTyrAspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaHisAsp
4622  CCAGAATACGACTTGGAGCTCATAACATCATGCTCCTCCAACGTGTCAGTCGCCCACGAC
      GGTCTTATGCTGAACCTCGAGTATTGTAGTACGAGGAGGTTGCACAGTCAGCGGGTGCTG
                                     ^
```

4637 SACI,

```
      GlyAlaGlyLysArgValTyrTyrLeuThrArgAspProThrThrProLeuAlaArgAla
4682  GGCGCTGGAAAGAGGGTCTACTACCTCACCCGTGACCCTACAACCCCCCTCGCGAGAGCT
      CCGCGACCTTTCTCCCAGATGATGGAGTGGGCACTGGGATGTTGGGGGGAGCGCTCTCGA
                                                           ^
```

4731 NRUI,

```
      AlaTrpGluThrAlaArgHisThrProValAsnSerTrpLeuGlyAsnIleIleMetPhe
4742  GCGTGGGAGACAGCAAGACACACTCCAGTCAATTCCTGGCTAGGCAACATAATCATGTTT
      CGCACCCTCTGTCGTTCTGTGTGAGGTCAGTTAAGGACCGATCCGTTGTATTAGTACAAA
```

```
      AlaProThrLeuTrpAlaArgMetIleLeuMetThrHisPhePheSerValLeuIleAla
4802  GCCCCCACACTGTGGGCGAGGATGATACTGATGACCCATTTCTTTAGCGTCCTTATAGCC
      CGGGGGTGTGACACCCGCTCCTACTATGACTACTGGGTAAAGAAATCGCAGGAATATCGG
             ^^
```

4806 PFLM1, 4807 DRA3,

```
      ArgAspGlnLeuGluGlnAlaLeuAspCysGluIleTyrGlyAlaCysTyrSerIleGlu
```

FIGURE 22 - Page 9

4862 AGGGACCAGCTTGAACAGGCCCTCGATTGCGAGATCTACGGGGCCTGCTACTCCATAGAA
     TCCCTGGTCGAACTTGTCCGGGAGCTAACGCTCTAGATGCCCCGGACGATGAGGTATCTT
                                        ^

     4893 BGL2,

     ProLeuAspLeuProProIleIleGlnArgLeuHisGlyLeuSerAlaPheSerLeuHis
4922 CCACTGGATCTACCTCCAATCATTCAAAGACTCCATGGCCTCAGCGCATTTTCACTCCAC
     GGTGACCTAGATGGAGGTTAGTAAGTTTCTGAGGTACCGGAGTCGCGTAAAAGTGAGGTG
                                    ^

     4954 NCOI,

     SerTyrSerProGlyGluIleAsnArgValAlaAlaCysLeuArgLysLeuGlyValPro
4982 AGTTACTCTCCAGGTGAAATCAATAGGGTGGCCGCATGCCTCAGAAAACTTGGGGTACCG
     TCAATGAGAGGTCCACTTTAGTTATCCCACCGGCGTACGGAGTCTTTTGAACCCCATGGC
                                   ^   ^

     5015 SPHI, 5035 KPNI,

     ProLeuArgAlaTrpArgHisArgAlaArgSerValArgAlaArgLeuLeuAlaArgGly
5042 CCCTTGCGAGCTTGGAGACACCGGGCCCGGAGCGTCCGCGCTAGGCTTCTGGCCAGAGGA
     GGGAACGCTCGAACCTCTGTGGCCCGGGCCTCGCAGGCGCGATCCGAAGACCGGTCTCCT
                        ^   ^

     5064 APAI, 5091 BALI,

     GlyArgAlaAlaIleCysGlyLysTyrLeuPheAsnTrpAlaValArgThrLysLeuLys
5102 GGCAGGGCTGCCATATGTGGCAAGTACCTCTTCAACTGGGCAGTAAGAACAAAGCTCAAA
     CCGTCCCGACGGTATACACCGTTCATGGAGAAGTTGACCCGTCATTCTTGTTTCGAGTTT
                    ^

     5113 NDEI,

     LeuThrProIleAlaAlaAlaGlyGlnLeuAspLeuSerGlyTrpPheThrAlaGlyTyr
5162 CTCACTCCAATAGCGGCCGCTGGCCAGCTGGACTTGTCCGGCTGGTTCACGGCTGGCTAC
     GAGTGAGGTTATCGCCGGCGACCGGTCGACCTGAACAGGCCGACCAAGTGCCGACCGATG
                    ^^   ^   ^

     5174 NOTI, 5175 EAG1 XMA3, 5182 BALI, 5186 PVU2,

     SerGlyGlyAspIleTyrHisSerValSerHisAlaArgProArgTrpIleTrpPheCys
5222 AGCGGGGGAGACATTTATCACAGCGTGTCTCATGCCCGGCCCCGCTGGATCTGGTTTTGC
     TCGCCCCCTCTGTAAATAGTGTCGCACAGAGTACGGGCCGGGGCGACCTAGACCAAAACG
                     ^

     5240 DRA3,

     LeuLeuLeuLeuAlaAlaGlyValGlyIleTyrLeuLeuProAsnArgMetSerThrAsn
5282 CTACTCCTGCTTGCTGCAGGGGTAGGCATCTACCTCCTCCCCAACCGAATGAGCACGAAT
     GATGAGGACGAACGACGTCCCCATCCGTAGATGGAGGAGGGGTTGGCTTACTCGTGCTTA
                    ^

     5295 PSTI,

     ProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGlnAspValLysPhe
5342 CCTAAACCTCAAAGAAAGACCAAACGTAACACCAACCGGCGGCCGCAGGACGTCAAGTTC
     GGATTTGGAGTTTCTTTCTGGTTTGCATTGTGGTTGGCCGCCGGCGTCCTGCAGTTCAAG
                                    ^^   ^   ^

     5380 NOTI, 5381 EAG1 XMA3, 5390 AAT2, 5401 SMAI XMAI,

     ProGlyGlyGlyGlnIleValGlyGlyValTyrLeuLeuProArgArgGlyProArgLeu
5402 CCGGGTGGCGGTCAGATCGTTGGTGGAGTTTACTTGTTGCCGCGCAGGGGCCCTAGATTG
     GGCCCACCGCCAGTCTAGCAACCACCTCAAATGAACAACGGCGCGTCCCCGGGATCTAAC
                                        ^

FIGURE 22 - Page 10

5449 APAI,

```
     GlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGlyArgArgGlnPro
5462 GGTGTGCGCGCGACGAGAAAGACTTCCGAGCGGTCGCAACCTCGAGGTAGACGTCAGCCT
     CCACACGCGCGCTGCTCTTTCTGAAGGCTCGCCAGCGTTGGAGCTCCATCTGCAGTCGGA
         ^              ^                    ^            ^
```

5467 BSSH2, 5478 XMNI, 5502 XHOI, 5511 AAT2,

```
     IleProLysAlaArgArgProGluGlyArgThrTrpAlaGlnProGlyTyrProTrpPro
5522 ATCCCCAAGGCTCGTCGGCCCGAGGGCAGGACCTGGGCTCAGCCCGGGTACCCTTGGCCC
     TAGGGGTTCCGAGCAGCCGGGCTCCCGTCCTGGACCCGAGTCGGGCCCATGGGAACCGGG
                          ^            ^      ^    ^
```

5548 ALWN1, 5558 ESP1, 5564 SMAI XMAI, 5568 KPNI,

```
     LeuTyrGlyAsnGluGlyCysGlyTrpAlaGlyTrpLeuLeuSerProArgGlySerArg
5582 CTCTATGGCAATGAGGGCTGCGGGTGGGCGGGATGGCTCCTGTCTCCCCGTGGCTCTCGG
     GAGATACCGTTACTCCCGACGCCCACCCGCCCTACCGAGGACAGAGGGGCACCGAGAGCC
```

```
     ProSerTrpGlyProThrAspProArgArgArgSerArgAsnLeuGlyLysValIleAsp
5642 CCTAGCTGGGGCCCCACAGACCCCCGGCGTAGGTCGCGCAATTTGGGTAAGGTCATCGAT
     GGATCGACCCCGGGGTGTCTGGGGGCCGCATCCAGCGCGTTAAACCCATTCCAGTAGCTA
         ^                                              ^
```

5650 APAI, 5696 CLAI,

```
     ThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProLeuValGlyAlaProLeu
5702 ACCCTTACGTGCGGCTTCGCCGACCTCATGGGGTACATACCGCTCGTCGGCGCCCCTCTT
     TGGGAATGCACGCCGAAGCGGCTGGAGTACCCCATGTATGGCGAGCAGCCGCGGGGAGAA
         ^                                    ^        ^
```
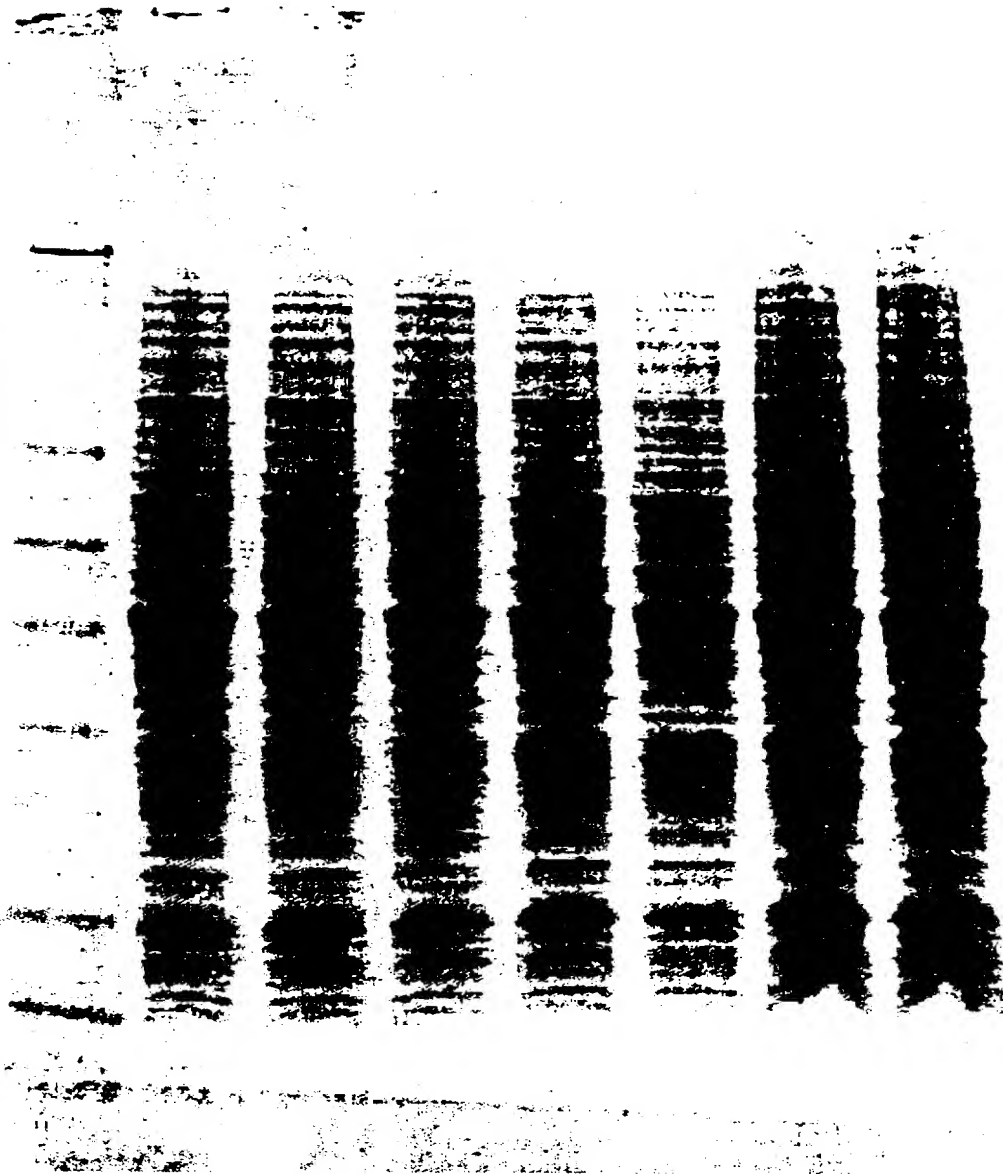
5724 HGIE2, 5750 KAS1 NARI, 5756 ECON1,

```
     GlyGlyAlaAlaArgAlaOC AM
5762 GGAGGCGCTGCCAGGGCCTAATAGTCGAC
     CCTCCGCGACGGTCCCGGATTATCAGCTG
                          ^
```

5785 SALI,

FIGURE 23